

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:25 ; Search time 39.0811 Seconds
(without alignments)
4076.547 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903
Sequence: 1 MALRGFCSADGSDPLMDMNV.....RSVAVAKKPKFSISPSLS 1927

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5 COMB.pdp: *
2: /cgn2_6/ptodata/1/iaa/6 COMB.pdp: *
3: /cgn2_6/ptodata/1/iaa/H_COMB.pdp: *
4: /cgn2_6/ptodata/1/iaa/PTIUS_COMB.pdp: *
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pdp: *
6: /cgn2_6/ptodata/1/iaa/backfillseq.pdp: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	79.4	1531	1 US-08-463-092B-4	Sequence 4, Appli
2	7860	79.4	1531	1 US-08-463-109A-4	Sequence 4, Appli
3	7860	79.4	1531	1 US-08-460-907B-4	Sequence 4, Appli
4	7860	79.4	1531	2 US-08-463-179A-4	Sequence 4, Appli
5	7860	79.4	1531	2 US-08-461-384B-4	Sequence 4, Appli
6	7860	79.4	1531	2 US-09-647-140B-19	Sequence 19, Appli
7	7849	79.3	1531	1 US-08-141-893-2	Sequence 2, Appli
8	7849	79.3	1531	1 US-08-463-092B-2	Sequence 2, Appli
9	7849	79.3	1531	1 US-08-463-109A-2	Sequence 2, Appli
10	7849	79.3	1531	1 US-08-460-907B-2	Sequence 2, Appli
11	7849	79.3	1531	2 US-08-463-179A-2	Sequence 2, Appli
12	7849	79.3	1531	2 US-08-461-384B-2	Sequence 2, Appli
13	7849	79.3	1531	2 US-08-407-207A-2	Sequence 2, Appli
14	7002.5	70.7	1528	1 US-08-463-092B-6	Sequence 6, Appli
15	7002.5	70.7	1528	1 US-08-463-109A-6	Sequence 6, Appli
16	7002.5	70.7	1528	1 US-08-460-907B-6	Sequence 6, Appli
17	7002.5	70.7	1528	2 US-08-463-179A-6	Sequence 6, Appli
18	7002.5	70.7	1528	2 US-08-461-384B-6	Sequence 6, Appli
19	4487.5	45.3	1527	2 US-09-647-140B-6	Sequence 33, Appli
20	4487.5	45.3	1527	2 US-09-647-140B-33	Sequence 33, Appli
21	3391.5	34.2	1503	2 US-09-647-140B-8	Sequence 8, Appli
22	3385.5	34.2	1503	2 US-09-792-616-3	Sequence 3, Appli
23	3218.5	32.5	1498	2 US-09-792-616-9	Sequence 9, Appli
24	2374.5	24.0	1621	2 US-08-972-927-3	Sequence 3, Appli
25	2349.5	23.7	1622	2 US-08-972-927-6	Sequence 6, Appli
26	2266.5	22.9	1325	2 US-09-647-140B-2	Sequence 2, Appli
27	2212.5	22.3	1464	2 US-10-012-896-1008	Sequence 1008, Ap

28	2195.5	22.2	1261	2 US-09-439-313-538	Sequence 538, App
29	2195.5	22.2	1261	2 US-09-636-215-538	Sequence 538, App
30	2195.5	22.2	1261	2 US-09-685-166A-538	Sequence 538, App
31	2195.5	22.2	1261	2 US-09-679-426-538	Sequence 538, App
32	2195.5	22.2	1261	2 US-09-759-143-538	Sequence 538, App
33	2195.5	22.2	1261	2 US-09-651-236-538	Sequence 538, App
34	2195.5	22.2	1261	2 US-09-657-279-538	Sequence 538, App
35	2195.5	22.2	1261	2 US-10-012-896-538	Sequence 538, App
36	2195.5	22.2	1261	2 US-10-012-896-1009	Sequence 1009, App
37	2118	21.4	1581	2 US-08-726-320-3	Sequence 3, Appli
38	2118	21.4	1581	2 US-09-208-716-3	Sequence 3, Appli
39	2115	21.4	1228	2 US-09-439-313-537	Sequence 537, App
40	2115	21.4	1228	2 US-09-636-215-537	Sequence 537, App
41	2115	21.4	1228	2 US-09-685-166A-537	Sequence 537, App
42	2115	21.4	1228	2 US-09-679-426-537	Sequence 537, App
43	2115	21.4	1228	2 US-09-759-143-537	Sequence 537, App
44	2115	21.4	1228	2 US-09-651-236-537	Sequence 537, App
45	2115	21.4	1228	2 US-09-657-279-537	Sequence 537, App

ALIGNMENTS

RESULT 1
US-08-463-092B-4
Sequence 4, Application US/08463092B
Patent No. 576680
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1531 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-463-092B-4

Query Match 79.4%; Score 7860; DB 1; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALRGFCSDGSDPLMDNNVTWNTSNPDPTKCFONTVLVWPCFYLMACPPFFLYLSRH 60
DB 1 MALRGFCSDGSDPLMDNNVTWNTSNPDPTKCFONTVLVWPCFYLMACPPFFLYLSRH 60
QY 61 DRGYIOMPLNKTATGALGFLIMIVCMADLFYSFMRSRSGIFLAPFVLSPILLGITLLA 120
DB 61 DRGYIOMPLNKTATGALGFLIMIVCMADLFYSFMRSRSGIFLAPFVLSPILLGITLLA 120
QY 61 DRGYIOMPLNKTATGALGFLIMIVCMADLFYSFMRSRSGIFLAPFVLSPILLGITLLA 120
DB 61 DRGYIOMPLNKTATGALGFLIMIVCMADLFYSFMRSRSGIFLAPFVLSPILLGITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFLVALVQALAILRSKIMTALKEDAQVDLFRDITFVYVS 180
DB 121 TFLIQLERRKGVSSGIMLTFLVALVQALAILRSKIMTALKEDAQVDLFRDITFVYVS 180
QY 121 TFLIQLERRKGVSSGIMLTFLVALVQALAILRSKIMTALKEDAQVDLFRDITFVYVS 180
DB 121 TFLIQLERRKGVSSGIMLTFLVALVQALAILRSKIMTALKEDAQVDLFRDITFVYVS 180
QY 181 LLLIQLVLSGCSGDSPLFSETHIDNPCESSASFLSRITFWMITGLIVRGROPLEGSD 240
DB 181 LLLIQLVLSGCSGDSPLFSETHIDNPCESSASFLSRITFWMITGLIVRGROPLEGSD 240
QY 181 LLLIQLVLSGCSGDSPLFSETHIDNPCESSASFLSRITFWMITGLIVRGROPLEGSD 240
DB 181 LLLIQLVLSGCSGDSPLFSETHIDNPCESSASFLSRITFWMITGLIVRGROPLEGSD 240
QY 241 LMSLNKEDTSROVVVLVKNMKKECAKTRKOPVKVYVSSKDPAPCKESSKVDANEVVAL 300
DB 241 LMSLNKEDTSROVVVLVKNMKKECAKTRKOPVKVYVSSKDPAPCKESSKVDANEVVAL 300
QY 241 LMSLNKEDTSROVVVLVKNMKKECAKTRKOPVKVYVSSKDPAPCKESSKVDANEVVAL 300
DB 241 LMSLNKEDTSROVVVLVKNMKKECAKTRKOPVKVYVSSKDPAPCKESSKVDANEVVAL 300
QY 301 IVKSPQKEMNPSLFVLYKTGPFYLMGFFPKAIDHLMFGSPQILKLIKRVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFVLYKTGPFYLMGFFPKAIDHLMFGSPQILKLIKRVNDTKAPD 360
QY 301 IVKSPQKEMNPSLFVLYKTGPFYLMGFFPKAIDHLMFGSPQILKLIKRVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFVLYKTGPFYLMGFFPKAIDHLMFGSPQILKLIKRVNDTKAPD 360
QY 361 WQGYTYTLVLTATLQTLVHQYHICFVSGMRIKTAVIGAVRKALVITNSARKSTV 420
DB 361 WQGYTYTLVLTATLQTLVHQYHICFVSGMRIKTAVIGAVRKALVITNSARKSTV 420
QY 361 WQGYTYTLVLTATLQTLVHQYHICFVSGMRIKTAVIGAVRKALVITNSARKSTV 420
DB 361 WQGYTYTLVLTATLQTLVHQYHICFVSGMRIKTAVIGAVRKALVITNSARKSTV 420
QY 421 GEIVNLMVDARFMDLATYINMISAPLOVTLALYLMNLSGPIVAGVAVMTVMVNVN 480
DB 421 GEIVNLMVDARFMDLATYINMISAPLOVTLALYLMNLSGPIVAGVAVMTVMVNVN 480
QY 421 GEIVNLMVDARFMDLATYINMISAPLOVTLALYLMNLSGPIVAGVAVMTVMVNVN 480
DB 421 GEIVNLMVDARFMDLATYINMISAPLOVTLALYLMNLSGPIVAGVAVMTVMVNVN 480
QY 481 AVMAKTKTYOVAHMKSKDNRIKLNNEILNGIKVLKLYAMELAFDKYLAIRQELKYLK 540
DB 481 AVMAKTKTYOVAHMKSKDNRIKLNNEILNGIKVLKLYAMELAFDKYLAIRQELKYLK 540
QY 481 AVMAKTKTYOVAHMKSKDNRIKLNNEILNGIKVLKLYAMELAFDKYLAIRQELKYLK 540
DB 481 AVMAKTKTYOVAHMKSKDNRIKLNNEILNGIKVLKLYAMELAFDKYLAIRQELKYLK 540
QY 541 KSAVLSAVGTPTWCTPPLVALCTFAVYVYIDENNILDAQTAFSVLAIPNIRPLNITLP 600
DB 541 KSAVLSAVGTPTWCTPPLVALCTFAVYVYIDENNILDAQTAFSVLAIPNIRPLNITLP 600
QY 541 KSAVLSAVGTPTWCTPPLVALCTFAVYVYIDENNILDAQTAFSVLAIPNIRPLNITLP 600
DB 541 KSAVLSAVGTPTWCTPPLVALCTFAVYVYIDENNILDAQTAFSVLAIPNIRPLNITLP 600
QY 601 MVISIVQASVSLKRLIFLSHEELPDSIERPRKDDGNTSITVRATFTWASDPT 660
DB 601 MVISIVQASVSLKRLIFLSHEELPDSIERPRKDDGNTSITVRATFTWASDPT 660
QY 601 MVISIVQASVSLKRLIFLSHEELPDSIERPRKDDGNTSITVRATFTWASDPT 660
DB 601 MVISIVQASVSLKRLIFLSHEELPDSIERPRKDDGNTSITVRATFTWASDPT 660
QY 661 LINGIFPSIPEGALVAVVGVCGKSSLSALLAEMDKYGHVAIKGSVAVYVQAMIOND 720
DB 661 LINGIFPSIPEGALVAVVGVCGKSSLSALLAEMDKYGHVAIKGSVAVYVQAMIOND 720
QY 661 LINGIFPSIPEGALVAVVGVCGKSSLSALLAEMDKYGHVAIKGSVAVYVQAMIOND 720
DB 661 LINGIFPSIPEGALVAVVGVCGKSSLSALLAEMDKYGHVAIKGSVAVYVQAMIOND 720
QY 721 SLRENILFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQKORVSLAR 780
DB 721 SLRENILFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQKORVSLAR 780
QY 721 SLRENILFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQKORVSLAR 780
DB 721 SLRENILFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQKORVSLAR 780
QY 781 AVYSNADYILPDDPLSADVAHGHKIFENVIGPKMKLKKRKLIVTHSGSYLPQVDVIV 840
DB 781 AVYSNADYILPDDPLSADVAHGHKIFENVIGPKMKLKKRKLIVTHSGSYLPQVDVIV 840
QY 781 AVYSNADYILPDDPLSADVAHGHKIFENVIGPKMKLKKRKLIVTHSGSYLPQVDVIV 840
DB 781 AVYSNADYILPDDPLSADVAHGHKIFENVIGPKMKLKKRKLIVTHSGSYLPQVDVIV 840
QY 841 MSGGKISEMGSYOEILLARDGAFAEFLRTASTEOBDAEENGVTGSGPGKEAKQEMGM 900
DB 841 MSGGKISEMGSYOEILLARDGAFAEFLRTASTEOBDAEENGVTGSGPGKEAKQEMGM 900
QY 841 MSGGKISEMGSYOEILLARDGAFAEFLRTASTEOBDAEENGVTGSGPGKEAKQEMGM 900
DB 841 MSGGKISEMGSYOEILLARDGAFAEFLRTASTEOBDAEENGVTGSGPGKEAKQEMGM 900
QY 901 LVTSAGKOLQROUSSSSSYSGDISRHNSLAELQKAEKKETWLMEDAKQOTGOYKL 960
DB 901 LVTSAGKOLQROUSSSSSYSGDISRHNSLAELQKAEKKETWLMEDAKQOTGOYKL 960

```

```

DB 901 LVTSAGKOLQROUSSSSSYSGDISRHNSLAELQKAEKKETWLMEDAKQOTGOYKL 960
QY 961 SVYWDYMKAIQLFISFSLFLFMCNHYASALASNYWLSLMTDDPIVNGTOHTKVRLSVYG 1020
DB 961 SVYWDYMKAIQLFISFSLFLFMCNHYASALASNYWLSLMTDDPIVNGTOHTKVRLSVYG 1020
QY 1021 ALGISGCIAVFGYSMAVSIIGIILASRCLHYDLHLSIIRSPMSFFERTPSGNLVNRSKEL 1080
DB 1021 ALGISGCIAVFGYSMAVSIIGIILASRCLHYDLHLSIIRSPMSFFERTPSGNLVNRSKEL 1080
QY 1021 ALGISGCIAVFGYSMAVSIIGIILASRCLHYDLHLSIIRSPMSFFERTPSGNLVNRSKEL 1080
DB 1021 ALGISGCIAVFGYSMAVSIIGIILASRCLHYDLHLSIIRSPMSFFERTPSGNLVNRSKEL 1080
QY 1081 DTYSNMLPEVYIKMMSGLFNVIGACIYIILATPIAIIIPBLGIYFFVQRPVYASSROL 1140
DB 1081 DTYSNMLPEVYIKMMSGLFNVIGACIYIILATPIAIIIPBLGIYFFVQRPVYASSROL 1140
QY 1081 DTYSNMLPEVYIKMMSGLFNVIGACIYIILATPIAIIIPBLGIYFFVQRPVYASSROL 1140
DB 1081 DTYSNMLPEVYIKMMSGLFNVIGACIYIILATPIAIIIPBLGIYFFVQRPVYASSROL 1140
QY 1141 KRLSVSRSPYSHFNFTLLGVSYIRAFEOERTIHOSDLKVDENQAYPSIYANWMLA 1200
DB 1141 KRLSVSRSPYSHFNFTLLGVSYIRAFEOERTIHOSDLKVDENQAYPSIYANWMLA 1200
QY 1201 VRLECVNACIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNI VA 1260
DB 1201 VRLECVNACIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNI VA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSWPOYGRVFRVYCLRYREDLPVL RHINVTINGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSWPOYGRVFRVYCLRYREDLPVL RHINVTINGG 1320
QY 1261 VERLKEYSETEKAPMOIOETAPSSWPOYGRVFRVYCLRYREDLPVL RHINVTINGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSWPOYGRVFRVYCLRYREDLPVL RHINVTINGG 1320
QY 1321 EKVGI VGTGAGKSSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPDPVLF 1380
DB 1321 EKVGI VGTGAGKSSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPDPVLF 1380
QY 1381 SGSLRMNLDPPSQYSDERWTSLELAHLKOFVSLPKLDHECAGEBNTSVGQROLVCL 1440
DB 1381 SGSLRMNLDPPSQYSDERWTSLELAHLKOFVSLPKLDHECAGEBNTSVGQROLVCL 1440
QY 1441 ARALLRTKTLVDEAPRAVDLETFDDLIQSTIRQFEDCTVLTAAHRLNTIMDTRYIVL 1500
DB 1441 ARALLRTKTLVDEAPRAVDLETFDDLIQSTIRQFEDCTVLTAAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMAKXAGLV 1531
DB 1501 DKGEIOEGAPSDLLQORGLFYSMAKXAGLV 1531

```

RESULT 2
 US-08-462-109A-4
 ; Sequence 4, Application US/08462109A

Patent No. 5882875
 ; GENERAL INFORMATION:
 ; APPLICANT: Cole, Susan P.C.
 ; APPLICANT: Deeley, Roger G.
 ; TITLE OF INVENTION: METHODS FOR IDENTIFYING
 ; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: LAHIVE & COCKFIELD
 ; STREET: 60 State Street, suite 510
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA

ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/462,109A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/966,923
 ; FILING DATE: 27-OCT-1992
 ; APPLICATION NUMBER: 08/029,340

FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-4

Query Match 79.4%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pred. NO. 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDNMVNTMNTSNPDFTKCFONTLVWVPCFYLMACPEFFFLYSRH 60
DB 1 MALRGFCSADGSDPLMDNMVNTMNTSNPDFTKCFONTLVWVPCFYLMACPEFFFLYSRH 60
QY 61 DRGTYQMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLADVFLVSPTLGITTLLA 120
DB 61 DRGTYQMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLADVFLVSPTLGITTLLA 120
QY 121 TPLIQERRRGVSSGIMLTFMLVALYCALAIIASKIMTALKEAOVDLPDITFYYPFS 180
DB 121 TPLIQERRRGVSSGIMLTFMLVALYCALAIIASKIMTALKEAOVDLPDITFYYPFS 180
QY 121 TPLIQERRRGVSSGIMLTFMLVALYCALAIIASKIMTALKEAOVDLPDITFYYPFS 180
DB 121 TPLIQERRRGVSSGIMLTFMLVALYCALAIIASKIMTALKEAOVDLPDITFYYPFS 180
QY 181 LMLIQLVLSGSDSPLEFSETIHDNPNCPRESSASFLSRITPFWITGILVGYRQPLSGSD 240
DB 181 LMLIQLVLSGSDSPLEFSETIHDNPNCPRESSASFLSRITPFWITGILVGYRQPLSGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKKCAKTRKQPVVYVSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKKCAKTRKQPVVYVSSKDPAPKSSKVDANEVEAL 300
QY 241 LMSLNKEDTSEQVVPVLVKNWKKCAKTRKQPVVYVSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKKCAKTRKQPVVYVSSKDPAPKSSKVDANEVEAL 300
QY 301 IVSPQKKNPSPLEKVLKTFGPFYFMSFFPKAIDHMFSGPQILKILIKFVNDTAPD 360
DB 301 IVSPQKKNPSPLEKVLKTFGPFYFMSFFPKAIDHMFSGPQILKILIKFVNDTAPD 360
QY 301 IVSPQKKNPSPLEKVLKTFGPFYFMSFFPKAIDHMFSGPQILKILIKFVNDTAPD 360
DB 301 IVSPQKKNPSPLEKVLKTFGPFYFMSFFPKAIDHMFSGPQILKILIKFVNDTAPD 360
QY 361 MOGYFYVLLFVNACTQTLVLAHOYFHI CFSGMR IKTAVI GAVYRKALVITNSARKSSTV 420
DB 361 MOGYFYVLLFVNACTQTLVLAHOYFHI CFSGMR IKTAVI GAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMLNLPGLSVAGVAVWMLVPV 480
DB 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMLNLPGLSVAGVAVWMLVPV 480
QY 481 AYWAMKTKTYQVAMKSKDNRIKLMNEILINGIKYLKLYAMELAFKDYLAIRQBELVVK 540
DB 481 AYWAMKTKTYQVAMKSKDNRIKLMNEILINGIKYLKLYAMELAFKDYLAIRQBELVVK 540
QY 541 KSAVLSAVGTETWCPPELVALCTFAYVYITDENNIIIDACTAFSLAFNIIIRPPLIIP 600
DB 541 KSAVLSAVGTETWCPPELVALCTFAYVYITDENNIIIDACTAFSLAFNIIIRPPLIIP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGGINSITVYNAFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGGINSITVYNAFTWARSDDPT 660
QY 661 LNTGTFSPREGALVAVVGVGCGSSLSLALLAEMDVREGHVAKGSVAVYPOQAVI QND 720
DB 661 LNTGTFSPREGALVAVVGVGCGSSLSLALLAEMDVREGHVAKGSVAVYPOQAVI QND 720
QY 721 SLRENILFGQLEPEYRSVYIQAALLPDEILIPSGDRTEIGEGVNLSSGQKQVSLAR 780

DB 721 SLRENILFGQLEPEYRSVYIQAALLPDEILIPSGDRTEIGEGVNLSSGQKQVSLAR 780
QY 781 AVYSNADITYPDDPLSANDVAGHIFENYVIGPKMKXKTRILLVTHSMGYLPQVDVYIV 840
DB 781 AVYSNADITYPDDPLSANDVAGHIFENYVIGPKMKXKTRILLVTHSMGYLPQVDVYIV 840
QY 841 MSGKISEMGSYOELLARQAPAEPLRTVASTOEOAEENGTVGSGPKKAKOMNGM 900
DB 841 MSGKISEMGSYOELLARQAPAEPLRTVASTOEOAEENGTVGSGPKKAKOMNGM 900
QY 901 LVYDSACKQOLQROUSSSSYSYSDISRHNSTAELOKAEKKEETKLMLEADKAQOTGVKL 960
DB 901 LVYDSACKQOLQROUSSSSYSYSDISRHNSTAELOKAEKKEETKLMLEADKAQOTGVKL 960
QY 961 SVYWDYKKAIGLPISLSTFLMCHNHSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYKKAIGLPISLSTFLMCHNHSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVGYSMAYSIGIILASRCLHYDLHSITRSFMSFPERPSGNLVNREKEL 1080
DB 1021 ALGISOGIAVGYSMAYSIGIILASRCLHYDLHSITRSFMSFPERPSGNLVNREKEL 1080
QY 1081 DTVDSMIPVYIKMFMSLFENVIGACTIYIILATPIAAIIPPLGIYFFVORFYVASSROL 1140
DB 1081 DTVDSMIPVYIKMFMSLFENVIGACTIYIILATPIAAIIPPLGIYFFVORFYVASSROL 1140
QY 1141 KRLSVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDIKVDENQAVYPSIVANRWLA 1200
DB 1141 KRLSVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDIKVDENQAVYPSIVANRWLA 1200
QY 1201 VRLCVCNCIVLFAALFAVISRHSLSAGVLSVSYSLQYTTVLANMLVRRSSSEMETIVA 1260
DB 1201 VRLCVCNCIVLFAALFAVISRHSLSAGVLSVSYSLQYTTVLANMLVRRSSSEMETIVA 1260
QY 1261 VERLKESETEKAPMOIOETAPSSWPQVGRVEFRNYCARYBEDDFVLRHINVTNGG 1320
DB 1261 VERLKESETEKAPMOIOETAPSSWPQVGRVEFRNYCARYBEDDFVLRHINVTNGG 1320
QY 1321 EKVIGVGTGAGKSSLTGLFRINESAGEIIDGINIAKIGHDLRFKTIIPQBPVLF 1380
DB 1321 EKVIGVGTGAGKSSLTGLFRINESAGEIIDGINIAKIGHDLRFKTIIPQBPVLF 1380
QY 1381 SGSLRNLDPFSQYSDBEVWTSLELAHKDFVSALPDKLDHECAEGGENLSVQORQVLC 1440
DB 1381 SGSLRNLDPFSQYSDBEVWTSLELAHKDFVSALPDKLDHECAEGGENLSVQORQVLC 1440
QY 1441 ARALLKTKILVDEATPAVDLETDDLIQSTIRTOEBDCVLTIAHRLNTIMDYTRYIVL 1500
DB 1441 ARALLKTKILVDEATPAVDLETDDLIQSTIRTOEBDCVLTIAHRLNTIMDYTRYIVL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531

RESULT 3
US-08-460-907B-4
Sequence 4, Application US/08460907B
Patent No. 5891724
GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
RESISTANCE ON A CELL
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: PARTIQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-4

Query Match      79.4%; Score 7860; DB 1; Length 1531;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MALRGFCAGSDSDPLMDNNTWNTSNPDFTKCFONTVLWVPCFYLMACPFFYLYLSRH 60
Db 1 MALRGFCAGSDSDPLMDNNTWNTSNPDFTKCFONTVLWVPCFYLMACPFFYLYLSRH 60
Oy 61 DRGYIOMTPIINKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGITLLA 120
Db 61 DRGYIOMTPIINKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGITLLA 120
Oy 121 TFLIQLERRKVOSSGIMLTFWLVALVCAALAIIRSKINTALKEDQVNDLFRDITFYVFS 180
Db 121 TFLIQLERRKVOSSGIMLTFWLVALVCAALAIIRSKINTALKEDQVNDLFRDITFYVFS 180
Oy 121 TFLIQLERRKVOSSGIMLTFWLVALVCAALAIIRSKINTALKEDQVNDLFRDITFYVFS 180
Db 121 TFLIQLERRKVOSSGIMLTFWLVALVCAALAIIRSKINTALKEDQVNDLFRDITFYVFS 180
Oy 181 LLLIQLVSCPSDSPLFSETIHDNPPCESSASFLSRTTFMTWITGLIVRGYROPLEGSD 240
Db 181 LLLIQLVSCPSDSPLFSETIHDNPPCESSASFLSRTTFMTWITGLIVRGYROPLEGSD 240
Oy 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
Oy 301 IVKSPQKEMNSLFLVLYKTGPFYLMSPFFKAIHDLMMFGSPQILKLIKRVNDTKAPD 360
Db 301 IVKSPQKEMNSLFLVLYKTGPFYLMSPFFKAIHDLMMFGSPQILKLIKRVNDTKAPD 360
Oy 361 WQGFYTYTLFETACLOTFLVHOYFHICFVSGMRITKTAVIGAVYKALVITNSAKKSTV 420
Db 361 WQGFYTYTLFETACLOTFLVHOYFHICFVSGMRITKTAVIGAVYKALVITNSAKKSTV 420
Oy 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLMLNLGSPVLGVAVMVLMEVNL 480
```

```

|||||
421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLMLNLGSPVLGVAVMVLMEVNL 480
Oy 481 ANAMAKTKTYQVAHMSKDNRIKLMEILNGIKVLKLYAMELAFKDKVLAIRQBELKVLK 540
Db 481 ANAMAKTKTYQVAHMSKDNRIKLMEILNGIKVLKLYAMELAFKDKVLAIRQBELKVLK 540
Oy 541 KSAVLSAVGFTWCTPEFLVALCTFAVYVTIDENNIIDAQTAFLVALFNILRPLNLLP 600
Db 541 KSAVLSAVGFTWCTPEFLVALCTFAVYVTIDENNIIDAQTAFLVALFNILRPLNLLP 600
Oy 601 MVISIVQASVSLKRLRIFLSHELEPDSIERRPVQDGGTNSITVNAFTWARSDDPT 660
Db 601 MVISIVQASVSLKRLRIFLSHELEPDSIERRPVQDGGTNSITVNAFTWARSDDPT 660
Oy 661 LMGITFSIPGALVAVAVGVGCGKSSILSLALBEMDVKEGHVAIKGSVAVYPOQAWIOND 720
Db 661 LMGITFSIPGALVAVAVGVGCGKSSILSLALBEMDVKEGHVAIKGSVAVYPOQAWIOND 720
Oy 721 SLRENILFGQLEBPYRSVYIACALLPDLIELPSGDRTEIGEKVNLSCGQKQSVLSAR 780
Db 721 SLRENILFGQLEBPYRSVYIACALLPDLIELPSGDRTEIGEKVNLSCGQKQSVLSAR 780
Oy 781 AVYSNADIYFPDDPLSAVDAAVGHKIFPENYIGPKMLKNTKRIILYTHSMGYLPQVDVYIV 840
Db 781 AVYSNADIYFPDDPLSAVDAAVGHKIFPENYIGPKMLKNTKRIILYTHSMGYLPQVDVYIV 840
Oy 841 MSGGKISEMSYQBELARDGAFAEFLRTYASTEODAEENGVTGVSQPKAEKAKOMENG 900
Db 841 MSGGKISEMSYQBELARDGAFAEFLRTYASTEODAEENGVTGVSQPKAEKAKOMENG 900
Oy 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAEILOKAEAKKEETWKLMEADKAQTGVKL 960
Db 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAEILOKAEAKKEETWKLMEADKAQTGVKL 960
Oy 961 SVYWDYKKAIGLFISFLSTFLPMCNRHVSALASNTWLSLMTDDPVLVNGTOHTKRLSVYG 1020
Db 961 SVYWDYKKAIGLFISFLSTFLPMCNRHVSALASNTWLSLMTDDPVLVNGTOHTKRLSVYG 1020
Oy 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHYDLHSILRSPMSFFERTPSGNLVNRFSEKL 1080
Db 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHYDLHSILRSPMSFFERTPSGNLVNRFSEKL 1080
Oy 1081 DTVDMSIPEVILKMGMSLFENVIGACIVILLATPILAIIPPLGLIYFVORFYVASSROL 1140
Db 1081 DTVDMSIPEVILKMGMSLFENVIGACIVILLATPILAIIPPLGLIYFVORFYVASSROL 1140
Oy 1141 KRLSVSRSPYSHFNFTLLGVSYTRAFEOERTIHOSDLKVDENOKAYPSIVANFWLA 1200
Db 1141 KRLSVSRSPYSHFNFTLLGVSYTRAFEOERTIHOSDLKVDENOKAYPSIVANFWLA 1200
Oy 1201 VRLCEVNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYVNLVRRMSSEMETNIVA 1260
Db 1201 VRLCEVNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYVNLVRRMSSEMETNIVA 1260
Oy 1261 VERLKEYSETEKEAPMOIOETAPPSWPOVGRVFRNYCLTRYEDLDFVLRIHIVTINGG 1320
Db 1261 VERLKEYSETEKEAPMOIOETAPPSWPOVGRVFRNYCLTRYEDLDFVLRIHIVTINGG 1320
Oy 1321 EKVGI VRTGAKSSLTGLFRINESABGEIIDGINIAKIGLHDLRFKTTIIQODPVLF 1380
Db 1321 EKVGI VRTGAKSSLTGLFRINESABGEIIDGINIAKIGLHDLRFKTTIIQODPVLF 1380
Oy 1381 SGRIRMLNDPSQVSDSEWMTSLELAHLKOFVSLPKLTHECAGEENLSVGOROLVCL 1440
Db 1381 SGRIRMLNDPSQVSDSEWMTSLELAHLKOFVSLPKLTHECAGEENLSVGOROLVCL 1440
Oy 1441 ARALIRKTKILVDEATAAVDFETDDIIOGTIRFQFEDCVLTIAHRLNTIMDTRVIVL 1500
Db 1441 ARALIRKTKILVDEATAAVDFETDDIIOGTIRFQFEDCVLTIAHRLNTIMDTRVIVL 1500
Oy 1501 DKGEIOEYGAESDILQOGLFYNSMAKDAGLV 1531
```

Db 1501 DKGEIOEYGAPELDLQORGLFYNSAKDAGLV 1531

RESULT 4
US-08-463-179A-4
Sequence 4, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deconci, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-4

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCAGSDPLMDMNTWNTSNDFPKCFONTLVLVWPCFYIMACFPYFLYSRH 60
Db 1 MALRGFCAGSDPLMDMNTWNTSNDFPKCFONTLVLVWPCFYIMACFPYFLYSRH 60
QY 61 DRGVIQMTPLNKTKTALGFLIMTVCMADLFYSFWEBSRGIFLAPVFLVSPFLIGITTLA 120
Db 61 DRGVIQMTPLNKTKTALGFLIMTVCMADLFYSFWEBSRGIFLAPVFLVSPFLIGITTLA 120
QY 121 TPLIQERRRGVSSGIMLTFMLVALVCAALILRSKIMTALKEDAQVDLFRDITTFYVFS 180
Db 121 TPLIQERRRGVSSGIMLTFMLVALVCAALILRSKIMTALKEDAQVDLFRDITTFYVFS 180
QY 181 LLLIQLVLSGFSRDRSPFLFSETIHDNCPRESSASFLSRITFWMTGLIVNGYRPLEGSD 240
Db 181 LLLIQLVLSGFSRDRSPFLFSETIHDNCPRESSASFLSRITFWMTGLIVNGYRPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKKECAKTRKQPVKVYSSKDPQPKSSVVDANEVEAL 300

Db 241 LMSLNKEDTSEQVVPVLVKNWKKECAKTRKQPVKVYSSKDPQPKSSVVDANEVEAL 300
QY 301 IVKSPQKKNPSPFLKLYTKFGPYFLMSFFPKA1HDLMSGQILKLIKPNVDTPAPD 360
Db 301 IVKSPQKKNPSPFLKLYTKFGPYFLMSFFPKA1HDLMSGQILKLIKPNVDTPAPD 360
QY 361 WOGVFTVLLFWACIOTLVLHOFHIFVSGNRKTAIVGAYRKALVITNSARKSSTV 420
Db 361 WOGVFTVLLFWACIOTLVLHOFHIFVSGNRKTAIVGAYRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVIALYLMNLGSPVLAVGAVWMLVWPN 480
Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVIALYLMNLGSPVLAVGAVWMLVWPN 480
QY 481 AVAMKTKTYQVAMHMSKONRIKLMNEILNGIIVLKLYAMELAFKDKVLA1ROBELKVLK 540
Db 481 AVAMKTKTYQVAMHMSKONRIKLMNEILNGIIVLKLYAMELAFKDKVLA1ROBELKVLK 540
QY 541 KSAVLAVNGFTFWCTPFLVALCTFAVYVITDENNILDQTAFLVSLAFNILEPPLILP 600
Db 541 KSAVLAVNGFTFWCTPFLVALCTFAVYVITDENNILDQTAFLVSLAFNILEPPLILP 600
QY 601 MVISSIVQASVSLKRLRIFLSHBELEPDSIERRPVVDGGTNSITVRNATFTWARSDDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHBELEPDSIERRPVVDGGTNSITVRNATFTWARSDDPT 660
QY 661 LMGITFSIPGALVAVVGVGCGKSSLSALLAEKMRVGHVAKGSVAVVPOAWIOND 720
Db 661 LMGITFSIPGALVAVVGVGCGKSSLSALLAEKMRVGHVAKGSVAVVPOAWIOND 720
QY 721 SLRENILFGQLEBPYRSTYQACALLPDEILIPSGDRPTIGKGVNLSCGQQRVSLAR 780
Db 721 SLRENILFGQLEBPYRSTYQACALLPDEILIPSGDRPTIGKGVNLSCGQQRVSLAR 780
QY 781 AVYSNADIYLFDDPLSANDAVHVKHIFENVIGPKGMKNKTRILLYTHSMGYLPOVDYIY 840
Db 781 AVYSNADIYLFDDPLSANDAVHVKHIFENVIGPKGMKNKTRILLYTHSMGYLPOVDYIY 840
QY 841 MSGGKISEMGSYQELLARDGAFELRTYASTBOEDABENGTVGSGPKKAKOMENG 900
Db 841 MSGGKISEMGSYQELLARDGAFELRTYASTBOEDABENGTVGSGPKKAKOMENG 900
QY 901 LVTDSAGKQOROLSSSSSYSGDISRHNSTALQRAEAKKETWLTMEADKAQOTGVKL 960
Db 901 LVTDSAGKQOROLSSSSSYSGDISRHNSTALQRAEAKKETWLTMEADKAQOTGVKL 960
QY 961 SVYMDYMKAIGLFISFSLFPLFCNHYVSALASNYMLSLMTDDPIVNGTOEHTKRLSVYG 1020
Db 961 SVYMDYMKAIGLFISFSLFPLFCNHYVSALASNYMLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGILASRCLHVDLHSLIRSPWSPFERTPSGNLVNRPSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGILASRCLHVDLHSLIRSPWSPFERTPSGNLVNRPSEKL 1080
QY 1081 DTYDSMIPEVIKMFMSLENVIGACIVILLATPIAIIIPPLGLIYFVQRFVASSROL 1140
Db 1081 DTYDSMIPEVIKMFMSLENVIGACIVILLATPIAIIIPPLGLIYFVQRFVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLIGSVIRAFEOEFHOSDLKVDENOKAYYPSIVANRWLA 1200
Db 1141 KRLSVSRSPVYSHFNETLIGSVIRAFEOEFHOSDLKVDENOKAYYPSIVANRWLA 1200
QY 1201 VRLCEVNCIIVPAALFAVISRHSLSAGLVGSVSYSLQVTTYLVNMLVRNMSMETNIYA 1260
Db 1201 VRLCEVNCIIVPAALFAVISRHSLSAGLVGSVSYSLQVTTYLVNMLVRNMSMETNIYA 1260
QY 1261 VERLKEYSFTEKAPMOIOTETAPSSWPQVGRFERNYCLARREDLDVFLRHINVTINGG 1320
Db 1261 VERLKEYSFTEKAPMOIOTETAPSSWPQVGRFERNYCLARREDLDVFLRHINVTINGG 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINAKIGLHDLRFKTIIPQDPVLV 1380

Db 1321 EKVGIVGTGAGKSLTIGLFRINESAGEIIDDGINIAKIGLHDLRKFTITIIPODPVL 1380
Qy 1381 SGLSRMLNLPDPSQYSDSEVWTSLELAHLKDFVSALPDXLDEHCAEGGENTL SVGOROLVCL 1440
Db 1381 SGLSRMLNLPDPSQYSDSEVWTSLELAHLKDFVSALPDXLDEHCAEGGENTL SVGOROLVCL 1440
Qy 1441 ARALLRKTIIVDEATAVVDLETDLLQSTIRQFEDCTVLTIAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTIIVDEATAVVDLETDLLQSTIRQFEDCTVLTIAHRLNTIMDYTRIVL 1500
Qy 1501 DKGEIOEYGAPSDLLQORGLFYMAKDGALV 1531
Db 1501 DKGEIOEYGAPSDLLQORGLFYMAKDGALV 1531

RESULT 5
US-08-461-384B-4
Sequence 4, Application US/08461384B
Patent No. 6025473
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: PATEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,384B
FILING DATE: 05-JUN-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01547
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-384B-4

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADSDPLMDNMVNTSNPDFTKCFONTVLVWVPCFYLMACFPFYLYLSRH 60
Db 1 MALRGFCSADSDPLMDNMVNTSNPDFTKCFONTVLVWVPCFYLMACFPFYLYLSRH 60
Qy 61 DRGIQWTPLNKTKTALGFLIMVCMADLFYSFWRSGRTFLAPVFLVSPTLGITTLLA 120
Db 61 DRGIQWTPLNKTKTALGFLIMVCMADLFYSFWRSGRTFLAPVFLVSPTLGITTLLA 120

Db 61 DRGIQWTPLNKTKTALGFLIMVCMADLFYSFWRSGRTFLAPVFLVSPTLGITTLLA 120
Qy 121 TFLIOLERRKGVSSGIMLTFMWLVATVCAIAIIRSKIMTALKEBAQVDLFRDITFYVYFS 180
Db 121 TFLIOLERRKGVSSGIMLTFMWLVATVCAIAIIRSKIMTALKEBAQVDLFRDITFYVYFS 180
Qy 181 LLILQVLVSGFSDRSLPFSSETIHDNPPCPSSASFLSRTTFMWITGLIVNGYROPLEGSD 240
Db 181 LLILQVLVSGFSDRSLPFSSETIHDNPPCPSSASFLSRTTFMWITGLIVNGYROPLEGSD 240
Qy 241 LMSLNKEDTSEOVVPLVYKMKKECAKTRKOPKVVYSSSDPAPKSSSVVDANEVEAL 300
Db 241 LMSLNKEDTSEOVVPLVYKMKKECAKTRKOPKVVYSSSDPAPKSSSVVDANEVEAL 300
Qy 301 IVKSPQKMNPSLFKVLKYKTFGPFYFLMSFFPKAIHDLMPESGQILKLIKFVNDTKAPD 360
Db 301 IVKSPQKMNPSLFKVLKYKTFGPFYFLMSFFPKAIHDLMPESGQILKLIKFVNDTKAPD 360
Qy 361 WQGYFYTVLLPVDACLOTVLVHOYFHI CFVSGMKRIKTAIVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFYTVLLPVDACLOTVLVHOYFHI CFVSGMKRIKTAIVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAORPMDLATYINMWSAPLOVILALYLLMNLGSPVLAGVAVWVLMVNVN 480
Db 421 GEIYNLMSVDAORPMDLATYINMWSAPLOVILALYLLMNLGSPVLAGVAVWVLMVNVN 480
Qy 481 AVMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVTLAROBELKVLK 540
Db 481 AVMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVTLAROBELKVLK 540
Qy 541 KSAVLSAVGFTFWCTPFLVALCTPAVYVTIDENNIIADQATVSLAFNILRFPNIIIP 600
Db 541 KSAVLSAVGFTFWCTPFLVALCTPAVYVTIDENNIIADQATVSLAFNILRFPNIIIP 600
Qy 601 MVTSIVQASVSLKRLIFLSHELEPDSIERPVPKGGTNSITVNAFTTARSDPT 660
Db 601 MVTSIVQASVSLKRLIFLSHELEPDSIERPVPKGGTNSITVNAFTTARSDPT 660
Qy 661 LNCITFSIPGALVAVVGVCGGKSLSLALAEMLKVEGHVAIKGSVAVVPQAMIQND 720
Db 661 LNCITFSIPGALVAVVGVCGGKSLSLALAEMLKVEGHVAIKGSVAVVPQAMIQND 720
Qy 721 SLRENIIIFGCOLLEPPYRSVIOACALIPDLIELIPSGDRTIEGKGVNLSCGQKORVSLAR 780
Db 721 SLRENIIIFGCOLLEPPYRSVIOACALIPDLIELIPSGDRTIEGKGVNLSCGQKORVSLAR 780
Qy 781 AVYSNADIYLPDPLSAVDHAGKHIFENVIIGPKMLKNTRLIVTHSMGYLPQVDVIIV 840
Db 781 AVYSNADIYLPDPLSAVDHAGKHIFENVIIGPKMLKNTRLIVTHSMGYLPQVDVIIV 840
Qy 841 MSGGKISEMSYOELARDAFAEFLRTYASTEQOEDAENGVTGVSGPKEAKOMENG 900
Db 841 MSGGKISEMSYOELARDAFAEFLRTYASTEQOEDAENGVTGVSGPKEAKOMENG 900
Qy 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYABIQKAEKKEETWKLMEADKAQTGVKL 960
Db 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYABIQKAEKKEETWKLMEADKAQTGVKL 960
Qy 961 SVYWDYKKAIGLPSLSTFLFMCNHYSAASNYWSLWTDPPVNGTOHTKRLSVYG 1020
Db 961 SVYWDYKKAIGLPSLSTFLFMCNHYSAASNYWSLWTDPPVNGTOHTKRLSVYG 1020
Qy 1021 ALGISOGIAVFGYMAVSGIIGLASRCGLHYDLHSIIRSPMSFFERPPSGNLVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSGIIGLASRCGLHYDLHSIIRSPMSFFERPPSGNLVNRFSKEL 1080
Qy 1081 DTVDSMIPEVIKMFMSLFNVVIGACIVILLATPAAIIPPLGIYFVORFYVASSROL 1140
Db 1081 DTVDSMIPEVIKMFMSLFNVVIGACIVILLATPAAIIPPLGIYFVORFYVASSROL 1140
Qy 1141 KRLSVRSRSPYSHFNETLLGVSYTARFEQERFIHOSDLKVDENQAAVYPSIYANWLA 1200
Db 1141 KRLSVRSRSPYSHFNETLLGVSYTARFEQERFIHOSDLKVDENQAAVYPSIYANWLA 1200

QY 1201 VRLCEVGNCLVLPALFAVIRSRHSLSAGLVSVSYSHQVTTYLNMLVVRMSSEMETNIVA 1260
Db 1201 VRLCEVGNCLVLPALFAVIRSRHSLSAGLVSVSYSHQVTTYLNMLVVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDVLRHINVTINGG 1320
QY 1321 EKVIGVGTGAGKSLTGLFRINESAGEIIDGINAKIGLHDLRKKITITIQDPVLF 1380
Db 1321 EKVIGVGTGAGKSLTGLFRINESAGEIIDGINAKIGLHDLRKKITITIQDPVLF 1380
QY 1381 SGSLRMLNDPSSQXSDDEWTSLSLAHLKDFVSALPDKLDHCAEGENLSVGRQVCL 1440
Db 1381 SGSLRMLNDPSSQXSDDEWTSLSLAHLKDFVSALPDKLDHCAEGENLSVGRQVCL 1440
QY 1441 ARALLRKTILVLEDEATAVDLETDLDIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500
Db 1441 ARALLRKTILVLEDEATAVDLETDLDIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500
QY 1501 DKGEIOEYGAUSDLLQORGLFYMAKXAGLV 1531
Db 1501 DKGEIOEYGAUSDLLQORGLFYMAKXAGLV 1531

RESULT 6
US-09-647-140B-19
Sequence 19, Application US/09647140B
Patent No. 6803184
GENERAL INFORMATION:
APPLICANT: Fox Chase Cancer Center
APPLICANT: Krush, Gary D.
APPLICANT: Lee, Kun
APPLICANT: Belinsky, Martin G.
APPLICANT: Bain, Lisa J.
TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
TITLE OF INVENTION: Nucleic Acids and Methods of Use Thereof
FILE REFERENCE: FCCC 98-02
CURRENT APPLICATION NUMBER: US/09/647, 140B
CURRENT FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: PCT/US99/06644
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/079,759
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/095,153
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-647-140B-19

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVNTSNPDKCFONTLVVWPCFYLMACPFYFLYSRH 60
Db 1 MALRGFCSADGSDPLMDMNTVNTSNPDKCFONTLVVWPCFYLMACPFYFLYSRH 60
QY 61 DRGYIQMTPLNKTITAGFLIMVCMADLFYSFWERSRGIFLADVFLVSPFLIGITTLA 120
Db 61 DRGYIQMTPLNKTITAGFLIMVCMADLFYSFWERSRGIFLADVFLVSPFLIGITTLA 120
QY 121 TPLIQERRRGVSSGIMLFWLVALCALAILSKMTALKBAQVDLRDITFEYYS 180
Db 121 TPLIQERRRGVSSGIMLFWLVALCALAILSKMTALKBAQVDLRDITFEYYS 180
QY 181 LLLIQLVLSGFSRDSPLFSETIHDNPNCPRESSAFLSRITFWWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQLVLSGFSRDSPLFSETIHDNPNCPRESSAFLSRITFWWITGLIVRGYRQPLEGSD 240

Db 181 LLLIQLVLSGFSRDSPLFSETIHDNPNCPRESSAFLSRITFWWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSBOVVPVLYKMKKECATRQOPVAVVYSSSDPAPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSBOVVPVLYKMKKECATRQOPVAVVYSSSDPAPKSSKVDANEVEAL 300
QY 301 IVKSPQKEMNPSSLKVLVYKTFGPFYFLMSFFPKAIDHLMFSGPOILKLLIKFVNDTAPD 360
Db 301 IVKSPQKEMNPSSLKVLVYKTFGPFYFLMSFFPKAIDHLMFSGPOILKLLIKFVNDTAPD 360
QY 361 WQGYFTVLLFVTAQLOTVLHQYFHI CFVSGMR IKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTAQLOTVLHQYFHI CFVSGMR IKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQRMDLATYINMTWSPLOYITALLYLMLNGSVYLAGVAVMTLAMPVN 480
Db 421 GEIYNLMSVDAQRMDLATYINMTWSPLOYITALLYLMLNGSVYLAGVAVMTLAMPVN 480
QY 481 AVAMAKTKTYQVAMKSKDNRIKLMNEILINGIKYLYAMWELAKDKVLAIRBELKYLK 540
Db 481 AVAMAKTKTYQVAMKSKDNRIKLMNEILINGIKYLYAMWELAKDKVLAIRBELKYLK 540
QY 541 KSAVLSAVGTFTVCTPFLVALCTFAVYVITIDENNIIDAQTA FVSLAFNLRPLNLP 600
Db 541 KSAVLSAVGTFTVCTPFLVALCTFAVYVITIDENNIIDAQTA FVSLAFNLRPLNLP 600
QY 601 MVISSIVQASVSLKRLIFLSHELEBDSIERRPVKDGGGNSITVENATFTWARSDDPT 660
Db 601 MVISSIVQASVSLKRLIFLSHELEBDSIERRPVKDGGGNSITVENATFTWARSDDPT 660
QY 661 LNGITFSPGALVAVVGVGCGGSSLSALLAEMDVGEHVAKGSVAVYPOQAMOND 720
Db 661 LNGITFSPGALVAVVGVGCGGSSLSALLAEMDVGEHVAKGSVAVYPOQAMOND 720
QY 721 SLRENIIFGQLEBPYRSVIOACALLPDEILPSGRTEIGEXVMISGGOKRVSILAR 780
Db 721 SLRENIIFGQLEBPYRSVIOACALLPDEILPSGRTEIGEXVMISGGOKRVSILAR 780
QY 781 AVYSNADIYLPDDPLSAVDHVGKHIFENVIIGPKMLKNKTRILVTHSMSYLPQVDYIIV 840
Db 781 AVYSNADIYLPDDPLSAVDHVGKHIFENVIIGPKMLKNKTRILVTHSMSYLPQVDYIIV 840
QY 841 MSGGKISEMSGYELLARDGAPAEFLTYASTOGEODAEENGVTGSGPGEAQMENG 900
Db 841 MSGGKISEMSGYELLARDGAPAEFLTYASTOGEODAEENGVTGSGPGEAQMENG 900
QY 901 LVYDSAGKOROLSSSSYSGDISRHNSGTAELOKAEKKEFTWKMEADKAOTGOVKL 960
Db 901 LVYDSAGKOROLSSSSYSGDISRHNSGTAELOKAEKKEFTWKMEADKAOTGOVKL 960
QY 961 SVYWDYKAIQGLFISFLIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTYKRLSVYG 1020
Db 961 SVYWDYKAIQGLFISFLIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTYKRLSVYG 1020
QY 1021 ALGISQIAVGYSMNVISGILLASRCLAHYDLHSITRSMSFPERPSGNLVNRFKEL 1080
Db 1021 ALGISQIAVGYSMNVISGILLASRCLAHYDLHSITRSMSFPERPSGNLVNRFKEL 1080
QY 1081 DTYDSMIPVYIKMFMSLFNVIACIYIILATPPIAIIIPPLGIYFFVORFYVASSROL 1140
Db 1081 DTYDSMIPVYIKMFMSLFNVIACIYIILATPPIAIIIPPLGIYFFVORFYVASSROL 1140
QY 1141 KRLESVRSRSPYSHFNETLIGSVIRAFEBQERFIHQSDLKVDENQXVYPSIVANRWLA 1200
Db 1141 KRLESVRSRSPYSHFNETLIGSVIRAFEBQERFIHQSDLKVDENQXVYPSIVANRWLA 1200
QY 1201 VRLCEVGNCLVLPALFAVIRSRHSLSAGLVSVSYSHQVTTYLNMLVVRMSSEMETNIVA 1260
Db 1201 VRLCEVGNCLVLPALFAVIRSRHSLSAGLVSVSYSHQVTTYLNMLVVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDVLRHINVTINGG 1320

QY 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLDLREKTIIPQDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLDLREKTIIPQDPVLF 1380
QY 1381 SGSIKRNMLDPFSQYSDERWMTSLBLAHKDVSAIPDLDEHCAEGGENLSVGQOLVCL 1440
DB 1381 SGSIKRNMLDPFSQYSDERWMTSLBLAHKDVSAIPDLDEHCAEGGENLSVGQOLVCL 1440
QY 1441 ABALARKTKILVLEBATAVLETDLLOSTIRTOFEDCTVLTJAHRLNTIMDYTRVVL 1500
DB 1441 ABALARKTKILVLEBATAVLETDLLOSTIRTOFEDCTVLTJAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYMAKDAGLV 1531

RESULT 7
US-08-141-893-2
Sequence 2, Application US/08141893
Patent No. 5489519
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deele, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/141,893
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923; 08/029,340
FILING DATE: 27-OCT-1992; 8-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Decont, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-141-893-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCADSGDPLMDNNTWNTSNPDFTKCFQNTVLVWPCFYLMACFPFYLYLSRH 60
DB 1 MALRGFCADSGDPLMDNNTWNTSNPDFTKCFQNTVLVWPCFYLMACFPFYLYLSRH 60
QY 61 DRGYQMPPLNKTALGFLIMIVCMADLFYSFWERSGIFLAPVFLVSPITLGIITLLA 120
DB 61 DRGYQMPPLNKTALGFLIMIVCMADLFYSFWERSGIFLAPVFLVSPITLGIITLLA 120

QY 121 TFLIQERRKGVSSGIMLTFMWLVAVYCALAIIRSKMTLAKEDAVDLFRDITFYVFS 180
DB 121 TFLIQERRKGVSSGIMLTFMWLVAVYCALAIIRSKMTLAKEDAVDLFRDITFYVFS 180
QY 181 LLLIQVLSCFSPRSPLFSEITHDPNCPRESSAFISRIFFWMTGILVGYRQPLGSD 240
DB 181 LLLIQVLSCFSPRSPLFSEITHDPNCPRESSAFISRIFFWMTGILVGYRQPLGSD 240
QY 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKQPVYVYSSKDPAQPKESSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKQPVYVYSSKDPAQPKESSKVDANEVEAL 300
QY 301 IVSPQKEMNPSLFKVLKTFEGPYFLMSFFPKAIDHLMFSGPOLKILIKFVNDYAPD 360
DB 301 IVSPQKEMNPSLFKVLKTFEGPYFLMSFFPKAIDHLMFSGPOLKILIKFVNDYAPD 360
QY 361 WQGFYTVLLFVTAQCTLVLHOYFHI CFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTVLLFVTAQCTLVLHOYFHI CFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRMDLATYINMTWSAPLOVIALYLLMNLGFSVLAGVAVVWLVVFN 480
DB 421 GEIVNLMSVDAQRMDLATYINMTWSAPLOVIALYLLMNLGFSVLAGVAVVWLVVFN 480
QY 481 AVMAKTKTYQVAHMKSKNRIKLMNEILNGIKVLKYANBELAKDYLAIROBELKVLK 540
DB 481 AVMAKTKTYQVAHMKSKNRIKLMNEILNGIKVLKYANBELAKDYLAIROBELKVLK 540
QY 541 KSAVLASVGTPTWCPTPELVALCTPFAVYVITDENNILDQATAPFSLPLNLRPLNLR 600
DB 541 KSAVLASVGTPTWCPTPELVALCTPFAVYVITDENNILDQATAPFSLPLNLRPLNLR 600
QY 601 MVISIVQASVSLKRLRIFLSHELEBDSIERREPKDGGGINSITVNNATFTWARSDEPT 660
DB 601 MVISIVQASVSLKRLRIFLSHELEBDSIERREPKDGGGINSITVNNATFTWARSDEPT 660
QY 661 LMGITPESIPEGALVAVVGVQCGKSLLSALLAEMDVYEGHVAIKSVAVYVPOQAMTND 720
DB 661 LMGITPESIPEGALVAVVGVQCGKSLLSALLAEMDVYEGHVAIKSVAVYVPOQAMTND 720
QY 721 SLRENILFGCOLPEPPYRSVYQACALPDLIELPSGRTIEGEGVNLSSGQKQVSLAR 780
DB 721 SLRENILFGCOLPEPPYRSVYQACALPDLIELPSGRTIEGEGVNLSSGQKQVSLAR 780
QY 781 AVYNNADILEPDDPLSAVDAHVGKIFENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYNNADILEPDDPLSAVDAHVGKIFENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGKISEMSYQELARDAFAEFLRTYASTEQEODAEENGVTGVSQPKAEAKOMENG 900
DB 841 MSGKISEMSYQELARDAFAEFLRTYASTEQEODAEENGVTGVSQPKAEAKOMENG 900
QY 901 LVTHSAGKOLOROLSSSSSYSGDISRHNSYAEIOXAFAKKEFTWKLMEADKATQCYKL 960
DB 901 LVTHSAGKOLOROLSSSSSYSGDISRHNSYAEIOXAFAKKEFTWKLMEADKATQCYKL 960
QY 961 SVYWDYKAIQLFISFISIFLFMCNHSALASNYWLSMTDPIYNGTOEHTKRLSVYG 1020
DB 961 SVYWDYKAIQLFISFISIFLFMCNHSALASNYWLSMTDPIYNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLLHSILRSPSPFERTPSGNLVNRFSEL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLLHSILRSPSPFERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMIPEVVKMGMSLFNVIGACIYILATPIAAIIIPPLGLIYFFVQGFYVASSQOL 1140
DB 1081 DTVDSMIPEVVKMGMSLFNVIGACIYILATPIAAIIIPPLGLIYFFVQGFYVASSQOL 1140
QY 1141 KRLSEVSRSPPVYSHFNETLIGSVYIRAFEBEORFIHOSDLKVDENOKAYYSIVANRMLA 1200
DB 1141 KRLSEVSRSPPVYSHFNETLIGSVYIRAFEBEORFIHOSDLKVDENOKAYYSIVANRMLA 1200
QY 1201 VRLBCVGNICIVLFAALPAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260


```

Db 1201 VRLECVGNCIVLFAALFAVIRSHSLAGVLGVSVSYLQVTTYLMVLVMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKAAWQIOETAPPSNPQGVVERVANCYARRELDVLRHINTVING 1320
Db 1261 VERLKEVSETEKAAWQIOETAPPSNPQGVVERVANCYARRELDVLRHINTVING 1320
Qy 1321 EKVGVGTGAGKSLTGLPRINSAAGEIITIDGINIAKIGLHDKRKITITIPDDPLVF 1380
Db 1321 EKVGVGTGAGKSLTGLPRINSAAGEIITIDGINIAKIGLHDKRKITITIPDDPLVF 1380
Qy 1381 SGSLRMNDPFSQYSDEEWTSLBLAHLKDFVSALPDLDHCAAGENLSVGQRYVCL 1440
Db 1381 SGSLRMNDPFSQYSDEEWTSLBLAHLKDFVSALPDLDHCAAGENLSVGQRYVCL 1440
Qy 1441 ARALLRKTKIIVLDEATAVDLETDLLIQSTIRIQFEDCVTLTAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTKIIVLDEATAVDLETDLLIQSTIRIQFEDCVTLTAHRLNTIMDYTRIVL 1500
Qy 1501 DKGEIOEGAPSDLLQQRGLFYSMAKXAGLV 1531
Db 1501 DKGEIOEGAPSDLLQQRGLFYSMAKXAGLV 1531

RESULT 8
US-08-463-092B-2
Sequence 2, Application US/08463092B
Patent No. 5766880
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463, 092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966, 923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029, 340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141, 893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407, 207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6653

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-092B-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRFGSADSDPLMDMNTVMTNSDPFTCPONTLVWVPCRYLACFPFFLYLSRH 60
Db 1 MALRFGSADSDPLMDMNTVMTNSDPFTCPONTLVWVPCRYLACFPFFLYLSRH 60
Qy 61 DRGYIOMTPLNKTALGFLMIYCMADLFYSFERSRGIFLAVFVLSPTLGIITLLA 120
Db 61 DRGYIOMTPLNKTALGFLMIYCMADLFYSFERSRGIFLAVFVLSPTLGIITLLA 120
Qy 121 TFLIQLERRKGVSSGIMLTFWLVVALYCALAIRSKIMTALKEDAOVDLFFDITFYVES 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVVALYCALAIRSKIMTALKEDAOVDLFFDITFYVES 180
Qy 181 LLLIQVLSCPSDPSLSEETIHDNPNCPRESSASFSLRTFWMITGLIVRGYRQPLESSD 240
Db 181 LLLIQVLSCPSDPSLSEETIHDNPNCPRESSASFSLRTFWMITGLIVRGYRQPLESSD 240
Qy 241 LMSLNKEDTSEQVVPVVKWKKKCAKTRKQPVKYVSSKDPAPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVVKWKKKCAKTRKQPVKYVSSKDPAPKSSKVDANEVEAL 300
Qy 301 IVKSPQKEMPSLEFKVLYKTFGPVLSFFPKAIDHLMFSGPOLIKLLIKFVNDTRAPD 360
Db 301 IVKSPQKEMPSLEFKVLYKTFGPVLSFFPKAIDHLMFSGPOLIKLLIKFVNDTRAPD 360
Qy 361 WQGYFYTVLLFVTAQCTLVLAHQYFHCIFSQGMKIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFYTVLLFVTAQCTLVLAHQYFHCIFSQGMKIKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMLNGPSVLAGVAVMLAMPVN 480
Db 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMLNGPSVLAGVAVMLAMPVN 480
Qy 481 AVMAKTKTYQVAMKSKDNRIKLMNBLINGIKVLYAMELAFKDVLAIRQELRYLK 540
Db 481 AVMAKTKTYQVAMKSKDNRIKLMNBLINGIKVLYAMELAFKDVLAIRQELRYLK 540
Qy 541 KSAVLAVGTFTWCTPFVALCTFAVYVITDENNILLDAQTAFFVSLAFNLRPPLILP 600
Db 541 KSAVLAVGTFTWCTPFVALCTFAVYVITDENNILLDAQTAFFVSLAFNLRPPLILP 600
Qy 601 MVISIVQASVSLKRLRIFLSHELEPDSIERPVKGGGNTSYVNAFTYMARSDPT 660
Db 601 MVISIVQASVSLKRLRIFLSHELEPDSIERPVKGGGNTSYVNAFTYMARSDPT 660
Qy 661 LMGITFSPGALVAVVGVGCGKSLLSALLAEMDVREBHVAIKGSVAVYPQAWI QND 720
Db 661 LMGITFSPGALVAVVGVGCGKSLLSALLAEMDVREBHVAIKGSVAVYPQAWI QND 720
Qy 721 SLRENILFGQLEPPYRSYIOACALLPDLIELPSGRTIEGKGVNLSSGQKRVSLAR 780
Db 721 SLRENILFGQLEPPYRSYIOACALLPDLIELPSGRTIEGKGVNLSSGQKRVSLAR 780
Qy 781 AVYSNADIYFDDPLSAVDAHVGHIPENYIGPKMLKNTRILIVTSMYSVLPQVDYIIV 840
Db 781 AVYSNADIYFDDPLSAVDAHVGHIPENYIGPKMLKNTRILIVTSMYSVLPQVDYIIV 840
Qy 841 MSGGKISEMSYQELLARDGAPAEFLTYASTREODAEENGVTGSGPGKEAKOMENGM 900
Db 841 MSGGKISEMSYQELLARDGAPAEFLTYASTREODAEENGVTGSGPGKEAKOMENGM 900
Qy 901 LVYDSACKQLQORQLSSSSSYSGDISRHNSGTAELOKAEAKKEEFWKLMEDAKQOTGVKL 960

```


Db 901 LVTSAGAOLOROSSSSSSSSGSDISRHNSTAELOKAKEETWKMEADAKOQGVKL 960
Qy 961 SVYDYMKAIGLIFSLIFLPMCHNVSALASNYLWLTMTDDPIYNGTOEHTKYVLSYVG 1020
Db 961 SVYDYMKAIGLIFSLIFLPMCHNVSALASNYLWLTMTDDPIYNGTOEHTKYVLSYVG 1020
Qy 1021 ALGISOGIAVFGYSMAVSIIGLILSRCLHVDLHLSILKSPMSFFERTPSGNIWNRFSKEL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGLILSRCLHVDLHLSILKSPMSFFERTPSGNIWNRFSKEL 1080
Qy 1081 DTUOSMIEVUKMPFGSLFNVIAGACIVILLATPIAIIIPPLGLTYFFVOGRFYVASSQOL 1140
Db 1081 DTUOSMIEVUKMPFGSLFNVIAGACIVILLATPIAIIIPPLGLTYFFVOGRFYVASSQOL 1140
Qy 1141 KRLESVSRSPYSHSHNETILGVSIVRAFEQERFIHQSDLKVDENQKAYPSIVANRWILA 1200
Db 1141 KRLESVSRSPYSHSHNETILGVSIVRAFEQERFIHQSDLKVDENQKAYPSIVANRWILA 1200
Qy 1201 VRLBCVNCIYLPALFAVISRHSLSAGLVGLSVSLSQVTTYLNWLVRMSSEMETNTIVA 1260
Db 1201 VRLBCVNCIYLPALFAVISRHSLSAGLVGLSVSLSQVTTYLNWLVRMSSEMETNTIVA 1260
Qy 1261 VERLKEVETREKAWOIOETRPSSWPQVGRVEFRNYCLARYEDLPVLRHINTVINGG 1320
Db 1261 VERLKEVETREKAWOIOETRPSSWPQVGRVEFRNYCLARYEDLPVLRHINTVINGG 1320
Qy 1321 EKVGVIGRTGAKSGLTGLFRINSAGEIITIDINIAKIGLHDLRFKITIIPODPVLF 1380
Db 1321 EKVGVIGRTGAKSGLTGLFRINSAGEIITIDINIAKIGLHDLRFKITIIPODPVLF 1380
Qy 1381 SGSLRMLNDPSPQSYDEEVTSLBLAHLKDFVSALPDKLDHCAEGENL SVGQRLVCL 1440
Db 1381 SGSLRMLNDPSPQSYDEEVTSLBLAHLKDFVSALPDKLDHCAEGENL SVGQRLVCL 1440
Qy 1441 ARALLRKRIIVLDATAVVDLETDDLIQSTIRTOPECTVLTIAHRLNTIMDYRIVVL 1500
Db 1441 ARALLRKRIIVLDATAVVDLETDDLIQSTIRTOPECTVLTIAHRLNTIMDYRIVVL 1500
Qy 1501 DKGEIOEYGA PSDLQORGLFYSMADAGLV 1531
Db 1501 DKGEIOEYGA PSDLQORGLFYSMADAGLV 1531

RESULT 9
US-08-462-109A-2
; Sequence 2, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Dealey, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992

APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCGADSDPLMDNVTWNTSNPDFYKCFONTVLVWPCTYLMACPEFFLYLSRH 60
Db 1 MALRGFCGADSDPLMDNVTWNTSNPDFYKCFONTVLVWPCTYLMACPEFFLYLSRH 60
Qy 61 DRGYIOMPLNKTATGALGFLIMTCWADLFYSPWERSGIFLAVFLVPSPTLLGITTLA 120
Db 61 DRGYIOMPLNKTATGALGFLIMTCWADLFYSPWERSGIFLAVFLVPSPTLLGITTLA 120
Qy 121 TFLIQLERRKVOSSGIMLTFMLVALYCALALILRSKIMTALKEDAOVDLFFRDIIFYVFS 180
Db 121 TFLIQLERRKVOSSGIMLTFMLVALYCALALILRSKIMTALKEDAOVDLFFRDIIFYVFS 180
Qy 181 LLLIQLVLSCFSDRSPLFSEFIHDNPNCPRESSASFLSHITFMWITGLIVRGYROPLEGSD 240
Db 181 LLLIQLVLSCFSDRSPLFSEFIHDNPNCPRESSASFLSHITFMWITGLIVRGYROPLEGSD 240
Qy 241 LMSLNKERTSQVVPVLYKMKKCECAKTRKQPVVYVYSSKDPAPKSSKYDANEVEVAL 300
Db 241 LMSLNKERTSQVVPVLYKMKKCECAKTRKQPVVYVYSSKDPAPKSSKYDANEVEVAL 300
Qy 301 IVKSPQKEMNSLFRVLYKTFGPFYLSFFPKAIIHDLMMFGSPOLIKLIFVNDTKAPD 360
Db 301 IVKSPQKEMNSLFRVLYKTFGPFYLSFFPKAIIHDLMMFGSPOLIKLIFVNDTKAPD 360
Qy 361 WQGYFYTLALFVTACLOTLVLAHQYFHHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFYTLALFVTACLOTLVLAHQYFHHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
Qy 421 GEIVNLMVDAQRFMDLATYINMTSAPLOVITLALYLMNLGSPVLAGVAVMLMVVN 480
Db 421 GEIVNLMVDAQRFMDLATYINMTSAPLOVITLALYLMNLGSPVLAGVAVMLMVVN 480
Qy 481 AVMAKTKTYOYAAHKSNDRIKLNNEILNGIKYVLYAMELAFODKYLAIROBELKYLK 540
Db 481 AVMAKTKTYOYAAHKSNDRIKLNNEILNGIKYVLYAMELAFODKYLAIROBELKYLK 540
Qy 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVITDENNIIDAQTAFFVALFNILRPLNLTLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVITDENNIIDAQTAFFVALFNILRPLNLTLP 600
Qy 601 MYSSIVASVSLKRLIFLSHEELPDSTIERRPYKDGCTNSTVVRNATFTWASDPT 660
Db 601 MYSSIVASVSLKRLIFLSHEELPDSTIERRPYKDGCTNSTVVRNATFTWASDPT 660
Qy 661 LNGITFESIPEGALVAVGVQCGKSLLSALLAEVDKYGHAIVKGSAAVYVQOAMIOND 720
Db 661 LNGITFESIPEGALVAVGVQCGKSLLSALLAEVDKYGHAIVKGSAAVYVQOAMIOND 720

QY 721 SLRENIIFGCOLLEBYRYSVIOACALLPDLLEIPSGDRTEIGKGVNLSSGOKORVSLAR 780
DB 721 SLRENIIFGCOLLEBYRYSVIOACALLPDLLEIPSGDRTEIGKGVNLSSGOKORVSLAR 780
QY 781 AVYSNADIIYLFDDPLSAVDAAVGHKIIFENVIGPKMKLNKTRILLVTHSMSTYLPQVDVITV 840
DB 781 AVYSNADIIYLFDDPLSAVDAAVGHKIIFENVIGPKMKLNKTRILLVTHSMSTYLPQVDVITV 840
QY 841 MSGKISMGSVOELLARDGAFAELRTYASTEOQDAEENGVCVSGPGEAKOMENGM 900
DB 841 MSGKISMGSVOELLARDGAFAELRTYASTEOQDAEENGVCVSGPGEAKOMENGM 900
QY 901 LVTSAGKOLOROLSSSSSSSGDISRHNSTAELOKAEKKEETKLMADKAOTGOYKL 960
DB 901 LVTSAGKOLOROLSSSSSSSGDISRHNSTAELOKAEKKEETKLMADKAOTGOYKL 960
QY 961 SVYWDYMKALIFISFLISIFLPMCNHVSALASNTWLSMTDDPIVNGTOEHKVKRLSYVG 1020
DB 961 SVYWDYMKALIFISFLISIFLPMCNHVSALASNTWLSMTDDPIVNGTOEHKVKRLSYVG 1020
QY 1021 ALGISOGIAVFGYGMNAVSGIIGILASRCLHNDLHLSILSPMSFPERPTSSGULNRFSEKL 1080
DB 1021 ALGISOGIAVFGYGMNAVSGIIGILASRCLHNDLHLSILSPMSFPERPTSSGULNRFSEKL 1080
QY 1081 DTVDSMIPEVIXMFMSGLFENVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSROL 1140
DB 1081 DTVDSMIPEVIXMFMSGLFENVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSROL 1140
QY 1141 KRLESVSRSPYVSHPNETLLGVSVYIARAEBOERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYVSHPNETLLGVSVYIARAEBOERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCTVLFPAALFVIVSRHSISAGLVLSVYSIQVTTYANLVKMSSEMETNIVA 1260
DB 1201 VRLCEVGNCTVLFPAALFVIVSRHSISAGLVLSVYSIQVTTYANLVKMSSEMETNIVA 1260
QY 1261 VERLKEVSETEKEAPMOIOETAPPSMPQVGRVFRANCCLAREBDLDPVLRHINVTNGS 1320
DB 1261 VERLKEVSETEKEAPMOIOETAPPSMPQVGRVFRANCCLAREBDLDPVLRHINVTNGS 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIDIINIAKIGLHDLREKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIDIINIAKIGLHDLREKTIIPQDPVLF 1380
QY 1381 SSGIRNMULDPSQYSDEBWTSLSLAHKDVYSALPKDLHCEKABGGENLSVGOPOLYCL 1440
DB 1381 SSGIRNMULDPSQYSDEBWTSLSLAHKDVYSALPKDLHCEKABGGENLSVGOPOLYCL 1440
QY 1441 ABALLRKTILVLEBATAVDLETTDDLIQSTIRTOFEDCTVLTJAHRLNTIMDTTRYIVL 1500
DB 1441 ABALLRKTILVLEBATAVDLETTDDLIQSTIRTOFEDCTVLTJAHRLNTIMDTTRYIVL 1500
QY 1501 DKGEIOEVGAPSDLLQORGLFYSNAKDAGLV 1531
DB 1501 DKGEIOEVGAPSDLLQORGLFYSNAKDAGLV 1531

RESULT 10
US-08-460-907B-2
Sequence 2, Application us/08460907B
Patent No. 5891724

GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P. C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
RESISTANCE ON A CELL
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA

ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: us/08/460, 907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966, 923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029, 340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141, 893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407, 207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTWNTSNDFTKCFONTLVVWPCFYLMACEPFYLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNTWNTSNDFTKCFONTLVVWPCFYLMACEPFYLYLSRH 60
QY 61 DRGIVQMTPLNKTKTALGFLIMTYVCADLYTSFMEBSRGIFLAPVFLVSTLLGITTLLA 120
DB 61 DRGIVQMTPLNKTKTALGFLIMTYVCADLYTSFMEBSRGIFLAPVFLVSTLLGITTLLA 120
QY 121 TPLIQLERRKGVSSGIMLTFWLVALVCAIAIIRSKITMTALKEDAQVDLFRDITFEYVYS 180
DB 121 TPLIQLERRKGVSSGIMLTFWLVALVCAIAIIRSKITMTALKEDAQVDLFRDITFEYVYS 180
QY 181 LLLIQLVLSGFSRSPPLFSETHIDPNCPRESSASFLSRTIFWMTGLIVAGYQPLEGSD 240
DB 181 LLLIQLVLSGFSRSPPLFSETHIDPNCPRESSASFLSRTIFWMTGLIVAGYQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVVKWKKCAKTRQPVKVVYSSQDPAQPKSSRYVDANEVEAL 300
DB 241 LMSLNKEDTSEOVVPLVVKWKKCAKTRQPVKVVYSSQDPAQPKSSRYVDANEVEAL 300
QY 301 IVKSPQKEMNPSEFKVLYKTFGYPFLMSFFPKAIHDMFSGQILEGLIKFVNDTAPD 360
DB 301 IVKSPQKEMNPSEFKVLYKTFGYPFLMSFFPKAIHDMFSGQILEGLIKFVNDTAPD 360
QY 361 WGGFYTVLLFVTRACIOTLVLHQYFHIQVSGMRITKTAIVGAYRKALVITNSARSSIV 420
DB 361 WGGFYTVLLFVTRACIOTLVLHQYFHIQVSGMRITKTAIVGAYRKALVITNSARSSIV 420

QY 421 GEIYNLMSVDAQRPMDLATYINMISAPLOVITLALYLMNLGSPVLGAVVMTLMPVN 480
DB 421 GEIYNLMSVDAQRPMDLATYINMISAPLOVITLALYLMNLGSPVLGAVVMTLMPVN 480
QY 481 AVAMAKTKTYOVAMHKSNDRIKLMNEILNGIKVLKLYAMELAFDKYLAIROEELKYLK 540
DB 481 AVAMAKTKTYOVAMHKSNDRIKLMNEILNGIKVLKLYAMELAFDKYLAIROEELKYLK 540
QY 541 KSAVLASAVGTFTWCTPELVALCTFAVYVITDENNILDQAOTAFVSLFNILRPLNLT 600
DB 541 KSAVLASAVGTFTWCTPELVALCTFAVYVITDENNILDQAOTAFVSLFNILRPLNLT 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPYKOGGCTNSITVRNATTTWASDPPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPYKOGGCTNSITVRNATTTWASDPPT 660
QY 661 LINGTFPSIPGALVAVVQVCGKSLSLALSALAEKDEKGVHAIKGAAYVPOQAMIOND 720
DB 661 LINGTFPSIPGALVAVVQVCGKSLSLALSALAEKDEKGVHAIKGAAYVPOQAMIOND 720
QY 721 SLRENIJFCQLEBEYRSVIOACALPDLBILPSGDRTEIGEKGVNLGGQKORVSLAR 780
DB 721 SLRENIJFCQLEBEYRSVIOACALPDLBILPSGDRTEIGEKGVNLGGQKORVSLAR 780
QY 781 AVYSNADTYLFPDDPLSANDAHYKHI FENVIGPKMLKXKTRILVTHSGNLYPOVDVILV 840
DB 781 AVYSNADTYLFPDDPLSANDAHYKHI FENVIGPKMLKXKTRILVTHSGNLYPOVDVILV 840
QY 841 MSGKISMGSYOEILARDGAPAEFLRTYASTBOCADBENGVTGVSQPGKAKOMENGM 900
DB 841 MSGKISMGSYOEILARDGAPAEFLRTYASTBOCADBENGVTGVSQPGKAKOMENGM 900
QY 901 LVTTDSAGKOLOROLSSSSSSSGDISRHHNSTALOKEAKKEETKLMKLEADKAOCTQVYL 960
DB 901 LVTTDSAGKOLOROLSSSSSSSGDISRHHNSTALOKEAKKEETKLMKLEADKAOCTQVYL 960
QY 961 SVYWDYMKAIQLFISFLSIFLPMCNHVSALASNYLSLMTDDPIVNGTOEHTKVLRYL 1020
DB 961 SVYWDYMKAIQLFISFLSIFLPMCNHVSALASNYLSLMTDDPIVNGTOEHTKVLRYL 1020
QY 1021 ALGTSQGIAPVGSMAVSIIGILASRCHVDLHSLRSPMSFFERTSGNLYNFSKEL 1080
DB 1021 ALGTSQGIAPVGSMAVSIIGILASRCHVDLHSLRSPMSFFERTSGNLYNFSKEL 1080
QY 1081 DTVSMIPEVIMKMFWSLFENVIGACTIVLLATPIAIIIPPLGIYFFVQRPYVASRQL 1140
DB 1081 DTVSMIPEVIMKMFWSLFENVIGACTIVLLATPIAIIIPPLGIYFFVQRPYVASRQL 1140
QY 1141 KRLESVSRSPYVSHNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYVPSIVANRWLA 1200
DB 1141 KRLESVSRSPYVSHNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYVPSIVANRWLA 1200
QY 1201 VRLGVGNCIYLFALPFAVISHSLSAGLVGSYSIQVTTYLMLVLRMSSEMTNIVA 1260
DB 1201 VRLGVGNCIYLFALPFAVISHSLSAGLVGSYSIQVTTYLMLVLRMSSEMTNIVA 1260
QY 1261 VERLKEVETEKAPMOIQETAPPSMPQVGRVFRNCLARREDLDFVLRINVTINGG 1320
DB 1261 VERLKEVETEKAPMOIQETAPPSMPQVGRVFRNCLARREDLDFVLRINVTINGG 1320
QY 1321 EKVGVIGRTGAGKSSLTGLFRINESABGEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
DB 1321 EKVGVIGRTGAGKSSLTGLFRINESABGEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
QY 1381 SCSLRMNIDPFSQYDEEVTWISLELAHLKDPYSALPKLDHSCAGGENTLSTGQOLVCL 1440
DB 1381 SCSLRMNIDPFSQYDEEVTWISLELAHLKDPYSALPKLDHSCAGGENTLSTGQOLVCL 1440
QY 1441 ARALIRKTKILVLDATAVDETDLIQSTIRTOPEDCVTLIAHRLNTIMDYTRVYL 1500
DB 1441 ARALIRKTKILVLDATAVDETDLIQSTIRTOPEDCVTLIAHRLNTIMDYTRVYL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531

DB 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531
RESULT 11
US-08-463-179A-2
; Sequence 2, Application US/08463179A
; Patent No. 6001563
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,179A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConli, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: POI-002CP8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-179A-2
Query Match 79.3%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALRGFCGADSDPFLMDNNVTWNTSNPDFTCFQNTVYLVWPCFYLMACPFYFLYLSRH 60
DB 1 MALRGFCGADSDPFLMDNNVTWNTSNPDFTCFQNTVYLVWPCFYLMACPFYFLYLSRH 60
QY 61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPTLGITTLA 120
DB 61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPTLGITTLA 120
QY 121 TFLIOLERRKVOSSGIMLTWLVALVCAALILRSKIMTALKEDAOVDLFRDITFYVYS 180
DB 121 TFLIOLERRKVOSSGIMLTWLVALVCAALILRSKIMTALKEDAOVDLFRDITFYVYS 180
QY 181 LLLIQLVLSCFSDSDSPFSETIHDNPPCESASFLSITFMWTGLIVRGROPLBGSD 240
DB 181 LLLIQLVLSCFSDSDSPFSETIHDNPPCESASFLSITFMWTGLIVRGROPLBGSD 240

QY 241 LMSLNKEDTSBQVVPVLVKNWKKCAKTRKQPVKVVYSSKDPAPQKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSBQVVPVLVKNWKKCAKTRKQPVKVVYSSKDPAPQKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLKVLYKTGPPFLMSFFFKAIHDLMEGSPQILKLLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLKVLYKTGPPFLMSFFFKAIHDLMEGSPQILKLLIFVNDTKAPD 360
QY 361 MGGFFVYTLFVTLCTOTLVHVOYFHCVFSGMRTKAVIGAVYRKALVITNSARKSSTV 420
DB 361 MGGFFVYTLFVTLCTOTLVHVOYFHCVFSGMRTKAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAQRFMIDLATYINMISAPLOVYIALMLGSLVLAAGVAVMLAMPVN 480
DB 421 GEIVNLMSVDAQRFMIDLATYINMISAPLOVYIALMLGSLVLAAGVAVMLAMPVN 480
QY 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDYLAIRQELKYLK 540
DB 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDYLAIRQELKYLK 540
QY 541 KSAVLAVGTFTWCTPPLVLCCTFAVYVTTIDENNILDAQAFVSLFNLRRPMLP 600
DB 541 KSAVLAVGTFTWCTPPLVLCCTFAVYVTTIDENNILDAQAFVSLFNLRRPMLP 600
QY 601 MVLSSIVQASVSLKRLIFLSHEELEPDSIERRPVKDGGGNSITVRNATFTMARSPPT 660
DB 601 MVLSSIVQASVSLKRLIFLSHEELEPDSIERRPVKDGGGNSITVRNATFTMARSPPT 660
QY 661 LMGITFESIPBGALVAVVGVQCGKSLSLALLAEMDYEGHVAIKGSVAVVPOQAMTOND 720
DB 661 LMGITFESIPBGALVAVVGVQCGKSLSLALLAEMDYEGHVAIKGSVAVVPOQAMTOND 720
QY 721 SLRENILFGCOLEBPYRSVQAQALLPDLEILPSGRTGEKGVLSSGQKRVSLAR 780
DB 721 SLRENILFGCOLEBPYRSVQAQALLPDLEILPSGRTGEKGVLSSGQKRVSLAR 780
QY 781 AVYSNADIYLFDDDLASVDAVHGKIFENVIGPKMLKNKRIIVTMSMSTLPQVDYIIV 840
DB 781 AVYSNADIYLFDDDLASVDAVHGKIFENVIGPKMLKNKRIIVTMSMSTLPQVDYIIV 840
QY 841 MSGGKISEMGSYOELLARDGAFELRTYASTEQDQAEENGVTGSGPKGKAKOMENGM 900
DB 841 MSGGKISEMGSYOELLARDGAFELRTYASTEQDQAEENGVTGSGPKGKAKOMENGM 900
QY 901 LMTDSAKKOLOROLSSSSSSSGDLSRHNSYAEIQAEBAKKEEYWKLEADKATQGVKL 960
DB 901 LMTDSAKKOLOROLSSSSSSSGDLSRHNSYAEIQAEBAKKEEYWKLEADKATQGVKL 960
QY 961 SVYWDYKKAIGLFIISFISIFLFCNHYSAASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYKKAIGLFIISFISIFLFCNHYSAASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGSMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGULVNRFSKEL 1080
DB 1021 ALGISOGIAVFGSMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGULVNRFSKEL 1080
QY 1081 DTYDSMTPEVITKMGSLFNVIACIYIILATPIAIIIPPLGIYFFVORFYASSROL 1140
DB 1081 DTYDSMTPEVITKMGSLFNVIACIYIILATPIAIIIPPLGIYFFVORFYASSROL 1140
QY 1141 KRLSVSRSPVSHFNETLGVSYIRAFEEQERFIHOSDLKVDENQKAYPSIYANMWA 1200
DB 1141 KRLSVSRSPVSHFNETLGVSYIRAFEEQERFIHOSDLKVDENQKAYPSIYANMWA 1200
QY 1201 VRLCEVNCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTYLNMVLRSSSEMETIIVA 1260
DB 1201 VRLCEVNCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTYLNMVLRSSSEMETIIVA 1260
QY 1261 VERLKESETEKEKPMOIOETAPSSSWPQVGRVAFNYCLARYRDLDPVLRHINVTNGG 1320
DB 1261 VERLKESETEKEKPMOIOETAPSSSWPQVGRVAFNYCLARYRDLDPVLRHINVTNGG 1320
QY 1321 EKVGIVGRTAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKLTIIIPQDPVLF 1380

DB 1321 EKVGIVGRTAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKLTIIIPQDPVLF 1380
QY 1381 SSGSLRMLDPPSOYSDEEWTSLLELAHLKDFVSLPKPKLHCEGGENISVGGROLYCL 1440
DB 1381 SSGSLRMLDPPSOYSDEEWTSLLELAHLKDFVSLPKPKLHCEGGENISVGGROLYCL 1440
QY 1441 ARALLRKTILVDEATRAVADLETDDLIQSTIRPQFEDCTVLTIAHRLNTIMDYTRYVL 1500
DB 1441 ARALLRKTILVDEATRAVADLETDDLIQSTIRPQFEDCTVLTIAHRLNTIMDYTRYVL 1500
QY 1501 DKGEIOEYGAPSDLLOQRGLFYSMKADAGLV 1531
DB 1501 DKGEIOEYGAPSDLLOQRGLFYSMKADAGLV 1531

RESULT 12
US-08-461-384B-2
; Sequence 2, Application US/08461384B
; Patent No. 6025473
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P. C.
; APPLICANT: Deele, Roger G.
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,384B
; FILING DATE: 05-JUN-95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Mlernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: 01547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-384B-2

Query Match 79.3%; Score 7849; DB 2; Length 1531;
Best local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGCSADGSDPLMDMNTWNTSNPDFTKCFQNTLVWVPCFYIMACFPYFLYLSRH 60
DB 1 MALRGCSADGSDPLMDMNTWNTSNPDFTKCFQNTLVWVPCFYIMACFPYFLYLSRH 60
QY 61 DRGYIQMTPLNKTALGLFLWVCAADLFYSFMBRSRGIFLAPVFLVSPFTLIGITTLA 120

```

Db      61 DRGIQMTPLNKRTALGFLLMIWCMADLFYSFWEKRGIFLAPFVLSPTLLITLLA 120
Qy      121 TFLQLERRKGVSSGIMLTFMLVALVCALAIIRSKITALKEDQVDFPDITFYVES 180
Db      121 TFLQLERRKGVSSGIMLTFMLVALVCALAIIRSKITALKEDQVDFPDITFYVES 180
Qy      181 LLLIQLVLSGSDSPLEFSETIHDNPSPRESSASFLSHITFMWITGLIVRGYROPLEGSD 240
Db      181 LLLIQLVLSGSDSPLEFSETIHDNPSPRESSASFLSHITFMWITGLIVRGYROPLEGSD 240
Qy      241 LMSLNKEDTSQVVPVLYVNMKKKCAKTRKQPVKVVYSSKDPAPQKSSKVDANEVVAL 300
Db      241 LMSLNKEDTSQVVPVLYVNMKKKCAKTRKQPVKVVYSSKDPAPQKSSKVDANEVVAL 300
Qy      301 IVKSPQKEMNPSLFVLYKTFGPFYFLMSFFPKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
Db      301 IVKSPQKEMNPSLFVLYKTFGPFYFLMSFFPKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
Qy      361 WQGFYTVLTFVTAQTLVLHOFPHICFVSGMRIKTAVIGAAYRKALVITNSARKSSTV 420
Db      361 WQGFYTVLTFVTAQTLVLHOFPHICFVSGMRIKTAVIGAAYRKALVITNSARKSSTV 420
Qy      421 GEIYNLMSVDQRFMDLATYINMISAPLOVILAYLMLNGPSVLGAVAVMVLMEVFN 480
Db      421 GEIYNLMSVDQRFMDLATYINMISAPLOVILAYLMLNGPSVLGAVAVMVLMEVFN 480
Qy      481 AVMAKTKTYOVANHKSNDNRKIKLMEILNGIKVLYLAMELAPKDKYLAIROEELKYLK 540
Db      481 AVMAKTKTYOVANHKSNDNRKIKLMEILNGIKVLYLAMELAPKDKYLAIROEELKYLK 540
Qy      541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNIIIDAQAFVSLAFNIIIRPPLNLP 600
Db      541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNIIIDAQAFVSLAFNIIIRPPLNLP 600
Qy      601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPYKOGGCTNSTVRAATTWARSDDPT 660
Db      601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPYKOGGCTNSTVRAATTWARSDDPT 660
Qy      661 LMGITFSPREGALVAVVGVQVCGKSLLSALLAEMDKVEGHVAKGSAAVYPOQAMIOND 720
Db      661 LMGITFSPREGALVAVVGVQVCGKSLLSALLAEMDKVEGHVAKGSAAVYPOQAMIOND 720
Qy      721 SLRENIIFGCOLLEBYRSVIOACALPLDLBITLPSGDRTEIGEKVNLSGGOKORVSLAR 780
Db      721 SLRENIIFGCOLLEBYRSVIOACALPLDLBITLPSGDRTEIGEKVNLSGGOKORVSLAR 780
Qy      781 AVYSNADIIYLFDDPLSAVDAAVGHKAIERNVIGPKMLKXKTRILVTHSMSTYLPQVDVILV 840
Db      781 AVYSNADIIYLFDDPLSAVDAAVGHKAIERNVIGPKMLKXKTRILVTHSMSTYLPQVDVILV 840
Qy      841 MSGGKISEMGSYOELLARDAFAEFLRTYASTBOQDAEBNGVGVSGPKGKAKOMENGM 900
Db      841 MSGGKISEMGSYOELLARDAFAEFLRTYASTBOQDAEBNGVGVSGPKGKAKOMENGM 900
Qy      901 LVYTSAGKOLOROLSSSSSSSGDISSRHNSTAELOKAEAKTEETKLMLEADQAOTGOYKL 960
Db      901 LVYTSAGKOLOROLSSSSSSSGDISSRHNSTAELOKAEAKTEETKLMLEADQAOTGOYKL 960
Qy      961 SVYNDYMKAIQLFISFLSIFLPMCHNVASALSNYMLSLMTDDPIYNGTQEHTRVLSYVG 1020
Db      961 SVYNDYMKAIQLFISFLSIFLPMCHNVASALSNYMLSLMTDDPIYNGTQEHTRVLSYVG 1020
Qy      1021 ALGTSQGIAPFGYSAVSIIGLILASRCLHVDLHSLSPMSFFERTSGNLYNFSKEL 1080
Db      1021 ALGTSQGIAPFGYSAVSIIGLILASRCLHVDLHSLSPMSFFERTSGNLYNFSKEL 1080
Qy      1081 DTVDSMIEVAKIMFMSLFENVGACIVILLATPILAIIPPLGILYFFQRFYVSSRQL 1140
Db      1081 DTVDSMIEVAKIMFMSLFENVGACIVILLATPILAIIPPLGILYFFQRFYVSSRQL 1140
Qy      1141 KRLSVSHSPYVSHENETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
Db      1141 KRLSVSHSPYVSHENETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200

```

```

Db      1141 KRLSVSHSPYVSHENETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
Qy      1201 VRLCEVGNCIYLFALFAVISRHSLSAGLVGLSVYSLOVTTYINMLVRMSSEMETNIVA 1260
Db      1201 VRLCEVGNCIYLFALFAVISRHSLSAGLVGLSVYSLOVTTYINMLVRMSSEMETNIVA 1260
Qy      1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVEFRNYCLRYRDDLPVLRHINVTINGG 1320
Db      1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVEFRNYCLRYRDDLPVLRHINVTINGG 1320
Qy      1321 EKVGIVGRTAGKSSLTGLFRINESAEGRIIIDGINIAKIGLHDLFPKTIIPQDPVLF 1380
Db      1321 EKVGIVGRTAGKSSLTGLFRINESAEGRIIIDGINIAKIGLHDLFPKTIIPQDPVLF 1380
Qy      1381 SGLSRMNLDPFSQYSDSEVWTSLELAHKDFVSAIPDKLDHECAEGENLSVGOROLVCL 1440
Db      1381 SGLSRMNLDPFSQYSDSEVWTSLELAHKDFVSAIPDKLDHECAEGENLSVGOROLVCL 1440
Qy      1441 ARALLRRTKIIIVDEATPAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
Db      1441 ARALLRRTKIIIVDEATPAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
Qy      1501 DKGIQIEYGAPSDLLQQRGLFYSMAXDAGLV 1531
Db      1501 DKGIQIEYGAPSDLLQQRGLFYSMAXDAGLV 1531

```

```

RESULT 13
US-08-407-207A-2
; Sequence 2, Application US/08407207A
; Patent No. 6063621
; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P. C.
; TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESS: PARTO RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,207A
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: 01512
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-407-207A-2

```

Query Match 79.3%; Score 7849; DB 2; Length 1531;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFSGADSDPLMDMNTVNTSNDPTKCFONTVLWVPCFYLAACFPPEFYLSRH 60
 DB 1 MALRGFSGADSDPLMDMNTVNTSNDPTKCFONTVLWVPCFYLAACFPPEFYLSRH 60

QY 61 DRGIIQMTPLNKTALGFLIMIVCADLFFSFWERSGIFLAIVFLVSPFLGITTLLA 120
 DB 61 DRGIIQMTPLNKTALGFLIMIVCADLFFSFWERSGIFLAIVFLVSPFLGITTLLA 120

QY 121 TFLIQLRRKGVSSGIMLTFMLVALVCAALILRSKIMTALKEAOVDLPFDITFEYFES 180
 DB 121 TFLIQLRRKGVSSGIMLTFMLVALVCAALILRSKIMTALKEAOVDLPFDITFEYFES 180

QY 181 LLLIQLVLSGFSRSPLESETIHDNPPCESSASFLSRITPEWITGLIVRGYRPLGSD 240
 DB 181 LLLIQLVLSGFSRSPLESETIHDNPPCESSASFLSRITPEWITGLIVRGYRPLGSD 240

QY 241 LMSLNKEDTSEQVIVLVKWKKCAKTRKOPVYVYVYSSKOPAPKSSKVDANEVYAL 300
 DB 241 LMSLNKEDTSEQVIVLVKWKKCAKTRKOPVYVYVYSSKOPAPKSSKVDANEVYAL 300

QY 301 IYKSPKEMNPSPLEKVLKTFGPFYFLMSFFPKAIDHLMFSGPQILKLIKFNVDTRAPD 360
 DB 301 IYKSPKEMNPSPLEKVLKTFGPFYFLMSFFPKAIDHLMFSGPQILKLIKFNVDTRAPD 360

QY 361 WOGFYTVLLFVTACLOTVLVHOYFHCFSVGNRIKTAIVIGAVRKALVITNSARKSSTV 420
 DB 361 WOGFYTVLLFVTACLOTVLVHOYFHCFSVGNRIKTAIVIGAVRKALVITNSARKSSTV 420

QY 421 GEIYNLMSVDAQRFMDIATYINMIWSAPLOYIALYLLMLNLSGVSILAGVAVMLVAVN 480
 DB 421 GEIYNLMSVDAQRFMDIATYINMIWSAPLOYIALYLLMLNLSGVSILAGVAVMLVAVN 480

QY 481 AYVAMKTKTYOVAMKSKDNRIKLMNELLINGIKYLKYAMELAKDVIAROSELVYK 540
 DB 481 AYVAMKTKTYOVAMKSKDNRIKLMNELLINGIKYLKYAMELAKDVIAROSELVYK 540

QY 541 KSAVLASVGTFTWCTPEPLVALCTFAVYVITDENNNIIDAQTAFFSLAFNLRPELILP 600
 DB 541 KSAVLASVGTFTWCTPEPLVALCTFAVYVITDENNNIIDAQTAFFSLAFNLRPELILP 600

QY 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGGTSITVNNATFTWARSDEPT 660
 DB 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGGTSITVNNATFTWARSDEPT 660

QY 661 LINGITFSPREGALVAVVGVGCCGKSLLSALLAEMDKVEGHVAKGSVAVVPOQAWI QND 720
 DB 661 LINGITFSPREGALVAVVGVGCCGKSLLSALLAEMDKVEGHVAKGSVAVVPOQAWI QND 720

QY 721 SLRENILFPGQLEBPYRSVYIACALPDLLEILPSGRTEIGEGVNLSSGQKORVSLAR 780
 DB 721 SLRENILFPGQLEBPYRSVYIACALPDLLEILPSGRTEIGEGVNLSSGQKORVSLAR 780

QY 781 AAVSNADIYLPDPLSAVDAHVGHIFENYIGPKMLKNKTRILVTHSMSYLPQVDYIIV 840
 DB 781 AAVSNADIYLPDPLSAVDAHVGHIFENYIGPKMLKNKTRILVTHSMSYLPQVDYIIV 840

QY 841 MSGGKISEMSYQBLARDAFAFLFTYASTEBEOAENGVGVSGPGEAKQOMENGM 900
 DB 841 MSGGKISEMSYQBLARDAFAFLFTYASTEBEOAENGVGVSGPGEAKQOMENGM 900

QY 901 LVTDASAKOLQRLSSSSSYSGDISRRHNSTAELOKAEAKKEETWKLMEADKQGTQVKL 960
 DB 901 LVTDASAKOLQRLSSSSSYSGDISRRHNSTAELOKAEAKKEETWKLMEADKQGTQVKL 960

QY 961 SVYWDYKAIAGLFTSFLSIFLPMGNVYSAASNYWLSLWTDPIVNGTOHTKVRISVYG 1020
 DB 961 SVYWDYKAIAGLFTSFLSIFLPMGNVYSAASNYWLSLWTDPIVNGTOHTKVRISVYG 1020

QY 1021 ALGISOGIAVFGYSAVYIGIILASRCLAHVLDLHLSILRSBMSFEERTPSGNLVNRFSEKL 1080
 DB 1021 ALGISOGIAVFGYSAVYIGIILASRCLAHVLDLHLSILRSBMSFEERTPSGNLVNRFSEKL 1080

QY 1081 DTVDSMIPEVYIKMFMSGLFNVIAGACIVILLATPIAIIIPPLGLIYEFVORFYVASSROL 1140
 DB 1081 DTVDSMIPEVYIKMFMSGLFNVIAGACIVILLATPIAIIIPPLGLIYEFVORFYVASSROL 1140

QY 1141 KRLSEVSRSVYSHFNFTLLGVSVIRAFEEQERFIHOSDLKDENQAYYPSIYANWLA 1200
 DB 1141 KRLSEVSRSVYSHFNFTLLGVSVIRAFEEQERFIHOSDLKDENQAYYPSIYANWLA 1200

QY 1201 VRLCEVNCIYLFALPAVIRSRSLAGVLGSYSLOYTTYANMLVRSSEMETNIVA 1260
 DB 1201 VRLCEVNCIYLFALPAVIRSRSLAGVLGSYSLOYTTYANMLVRSSEMETNIVA 1260

QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQYGRVEFRNYCLRYREDLDFVLRIHINTINGG 1320
 DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQYGRVEFRNYCLRYREDLDFVLRIHINTINGG 1320

QY 1321 EKVGIIVRTGAGKSSLTGLFRINESABGEIIIDGINIAKIGLHDLFKITIIIPQDPVLF 1380
 DB 1321 EKVGIIVRTGAGKSSLTGLFRINESABGEIIIDGINIAKIGLHDLFKITIIIPQDPVLF 1380

QY 1381 SGLSRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPDKLHECAGGENTSVGQROLVCL 1440
 DB 1381 SGLSRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPDKLHECAGGENTSVGQROLVCL 1440

QY 1441 ABALLRKTILVDEATAVADLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDTRVIVL 1500
 DB 1441 ABALLRKTILVDEATAVADLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDTRVIVL 1500

QY 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGIV 1531
 DB 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGIV 1531

RESULT 14
 US-08-463-092B-6
 ; Sequence 6, Application US/08463092B
 ; Patent No. 5766880
 ; GENERAL INFORMATION:
 ; APPLICANT: Cole, Susan P.C.
 ; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
 ; STREET: Queen's University at Kingston
 ; CITY: Kingston
 ; STATE: Ontario
 ; COUNTRY: CANADA
 ; ZIP: K7L 3N6
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/463,092B
 ; FILING DATE: 05-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/966,923
 ; FILING DATE: 27-OCT-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/029,340
 ; FILING DATE: 8-MAR-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/141,893

FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-092B-6

Query Match 70.7%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFCGADGSDPLMDNVTWNTSNPDFTKCFONTVLWVPCFYLMACFPFFYLYLSRH 60
DB 1 MALRSCFADGSDPLMDNVTWNTSNPDFTKCFONTVLWVPCFYLMACFPFFYLYLSRH 60
QY 61 DRGIQNTPLNKTALGFLMIYVCMADLFYSFWRSGQVLRAPVLVSPILLGITMLLA 120
DB 61 DRGIQNTPLNKTALGFLMIYVCMADLFYSFWRSGQVLRAPVLVSPILLGITMLLA 120
QY 121 TFLQLERKRVQSSGIMLTFLVALVCLALILRSKIMTALKEADQVDFPDIFPYVES 180
DB 121 TFLQLERKRVQSSGIMLTFLVALVCLALILRSKIMTALKEADQVDFPDIFPYVES 180
QY 121 TFLQLERKRVQSSGIMLTFLVALVCLALILRSKIMTALKEADQVDFPDIFPYVES 180
DB 121 TFLQLERKRVQSSGIMLTFLVALVCLALILRSKIMTALKEADQVDFPDIFPYVES 180
QY 181 LLLIQLVLSCFSDRSPLESETIHDPNCPRESSASFLSRTTWMITGLIVRGROPLEGSD 240
DB 181 LVLVOLVLSCFSDRSPLESETIHDPNCPRESSASFLSRTTWMITGLIVRGROPLEGSD 240
QY 241 LMSLNKETSQOVNVPVLYKMKKECATRKQPVKVVYS-SKDPAPQKSSKVDANEVEA 299
DB 241 LMSLNKETSQOVNVPVLYKMKKECATRKQPVKVVYS-SKDPAPQKSSKVDANEVEA 299
QY 241 LMSLNKETSQOVNVPVLYKMKKECATRKQPVKVVYS-SKDPAPQKSSKVDANEVEA 300
DB 241 LMSLNKETSQOVNVPVLYKMKKECATRKQPVKVVYS-SKDPAPQKSSKVDANEVEA 300
QY 300 LIVSPQKNSPLSKVLTKTGPYFLMSFPFKAIHDMFSGPOILKLIKPVNDTAP 359
DB 300 LIVSPQKNSPLSKVLTKTGPYFLMSFPFKAIHDMFSGPOILKLIKPVNDTAP 359
QY 301 LIVSPQKNSPLSKVLTKTGPYFLMSFPFKAIHDMFSGPOILKLIKPVNDTAP 360
DB 301 LIVSPQKNSPLSKVLTKTGPYFLMSFPFKAIHDMFSGPOILKLIKPVNDTAP 360
QY 360 DMGGFYTVLLFVTAQIOTLVLMHGFHICFVSGMRKITAIVIGAVYRKALVITNSARKSST 419
DB 360 DMGGFYTVLLFVTAQIOTLVLMHGFHICFVSGMRKITAIVIGAVYRKALVITNSARKSST 419
QY 361 DMGGFYTVLLFVTAQIOTLVLMHGFHICFVSGMRKITAIVIGAVYRKALVITNSARKSST 420
DB 361 DMGGFYTVLLFVTAQIOTLVLMHGFHICFVSGMRKITAIVIGAVYRKALVITNSARKSST 420
QY 420 VGEIVNLMSVDAQRPMDLATTINMIWASAPLOVITLALYMLNLGSPVLAQVAVVLMVPL 479
DB 420 VGEIVNLMSVDAQRPMDLATTINMIWASAPLOVITLALYMLNLGSPVLAQVAVVLMVPL 479
QY 421 VGEIVNLMSVDAQRPMDLATTINMIWASAPLOVITLALYMLNLGSPVLAQVAVVLMVPL 480
DB 421 VGEIVNLMSVDAQRPMDLATTINMIWASAPLOVITLALYMLNLGSPVLAQVAVVLMVPL 480
QY 480 NAWNAMKTKTYQVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAKDVLAIRBELKVL 539
DB 480 NAWNAMKTKTYQVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAKDVLAIRBELKVL 539
QY 481 NAWNAMKTKTYQVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAKDVLAIRBELKVL 540
DB 481 NAWNAMKTKTYQVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAKDVLAIRBELKVL 540
QY 540 KKSAYLSAVGTFTWCTPFLVALCTFAYVYVIDENNIIIDAQTAFFVSLAFNILEPPLIL 599
DB 540 KKSAYLSAVGTFTWCTPFLVALCTFAYVYVIDENNIIIDAQTAFFVSLAFNILEPPLIL 599
QY 541 KKSAYLSAVGTFTWCTPFLVALCTFAYVYVIDENNIIIDAQTAFFVSLAFNILEPPLIL 600
DB 541 KKSAYLSAVGTFTWCTPFLVALCTFAYVYVIDENNIIIDAQTAFFVSLAFNILEPPLIL 600
QY 600 PMVTSIVQASVSLKRLRIFLSHEELPDSIERRPVQGGGNTSTVNNATFTMARSDP 659
DB 600 PMVTSIVQASVSLKRLRIFLSHEELPDSIERRPVQGGGNTSTVNNATFTMARSDP 659
QY 601 PMVTSIVQASVSLKRLRIFLSHEELPDSIERRPVQGGGNTSTVNNATFTMARSDP 659
DB 601 PMVTSIVQASVSLKRLRIFLSHEELPDSIERRPVQGGGNTSTVNNATFTMARSDP 659
QY 660 TLNGITTSIPGALVAVVGVGGCGSSLLSALLAMDMKVEGVALKGSVAVVPOQAMION 719
DB 660 TLNGITTSIPGALVAVVGVGGCGSSLLSALLAMDMKVEGVALKGSVAVVPOQAMION 719

QY 720 DSIARENILFGCQLEBPYRSYIOACALLPDLEILPSSDRTEIGRKVNLSSGQKQRYSLA 779
DB 720 DSIARENILFGCQLEBPYRSYIOACALLPDLEILPSSDRTEIGRKVNLSSGQKQRYSLA 779
QY 780 RAVYSNADITLFDPLSAVDAHVGKHFENVIGPKGMLKNKTRTLVTHMSYLPQVDTI 839
DB 780 RAVYSNADITLFDPLSAVDAHVGKHFENVIGPKGMLKNKTRTLVTHMSYLPQVDTI 839
QY 840 VMSGKISMGYSYOELLARDGAFAPLRTYASTEODAEENGTVGSPGKEAKOMENG 899
DB 840 VMSGKISMGYSYOELLARDGAFAPLRTYASTEODAEENGTVGSPGKEAKOMENG 899
QY 840 VMSGKISMGYSYOELLARDGAFAPLRTYASTEODAEENGTVGSPGKEAKOMENG 896
DB 840 VMSGKISMGYSYOELLARDGAFAPLRTYASTEODAEENGTVGSPGKEAKOMENG 896
QY 900 MLVTDSAGKQOLROLSSSSSYSGDISHHNSTAELQALAKKETWKLMEADKQOTQYK 959
DB 900 MLVTDSAGKQOLROLSSSSSYSGDISHHNSTAELQALAKKETWKLMEADKQOTQYK 959
QY 960 LSVYNDYMKALIGLISFLSLFLPFCNHNVSALASNYWLSMTDD-PIYNGOEHKVLVS 1018
DB 960 LSVYNDYMKALIGLISFLSLFLPFCNHNVSALASNYWLSMTDD-PIYNGOEHKVLVS 1018
QY 966 LSVYNDYMKALIGLISFLSLFLPFCNHNVSALASNYWLSMTDD-PIYNGOEHKVLVS 1015
DB 966 LSVYNDYMKALIGLISFLSLFLPFCNHNVSALASNYWLSMTDD-PIYNGOEHKVLVS 1015
QY 1019 YGALGISQGIAPFCYSNAVSIQGIILASRCIHDVLSILNSPMSFFERTPSGNLVNRFSK 1078
DB 1019 YGALGISQGIAPFCYSNAVSIQGIILASRCIHDVLSILNSPMSFFERTPSGNLVNRFSK 1078
QY 1016 YGALGIIQGAALFQYSNAVSIQGIILASRCIHDVLSILNSPMSFFERTPSGNLVNRFSK 1075
DB 1016 YGALGIIQGAALFQYSNAVSIQGIILASRCIHDVLSILNSPMSFFERTPSGNLVNRFSK 1075
QY 1079 ELDTVDNMIBVYIMENGSFLPNVIGACTIVILATPAAIIPPLGLYFFQRYVASSR 1138
DB 1079 ELDTVDNMIBVYIMENGSFLPNVIGACTIVILATPAAIIPPLGLYFFQRYVASSR 1138
QY 1076 ELDTVDNMIBVYIMENGSFLPNVIGACTIVILATPAAIIPPLGLYFFQRYVASSR 1135
DB 1076 ELDTVDNMIBVYIMENGSFLPNVIGACTIVILATPAAIIPPLGLYFFQRYVASSR 1135
QY 1139 QLRKLESVRSPPVYSHNETLLGVSIRAFEBQERFHOSDLKVDENQKAYPSIVANRW 1198
DB 1139 QLRKLESVRSPPVYSHNETLLGVSIRAFEBQERFHOSDLKVDENQKAYPSIVANRW 1198
QY 1136 QLRKLESVRSPPVYSHNETLLGVSIRAFEBQERFHOSDLKVDENQKAYPSIVANRW 1195
DB 1136 QLRKLESVRSPPVYSHNETLLGVSIRAFEBQERFHOSDLKVDENQKAYPSIVANRW 1195
QY 1199 LAVALECVGNCTIVLFAALFVIRSHSLASGLVGLSVSLQVTTYLNLVMSSEMETNI 1258
DB 1199 LAVALECVGNCTIVLFAALFVIRSHSLASGLVGLSVSLQVTTYLNLVMSSEMETNI 1258
QY 1196 LAVALECVGNCTIVLFAALFVIRSHSLASGLVGLSVSLQVTTYLNLVMSSEMETNI 1255
DB 1196 LAVALECVGNCTIVLFAALFVIRSHSLASGLVGLSVSLQVTTYLNLVMSSEMETNI 1255
QY 1259 VAVRLKXSETEKEAPMOQIETAPPSGWPQVGEFRNRYCLARYEDLDVLAHINNTIN 1318
DB 1259 VAVRLKXSETEKEAPMOQIETAPPSGWPQVGEFRNRYCLARYEDLDVLAHINNTIN 1318
QY 1256 VAVRLKXSETEKEAPMOQIETAPPSGWPQVGEFRNRYCLARYEDLDVLAHINNTIN 1315
DB 1256 VAVRLKXSETEKEAPMOQIETAPPSGWPQVGEFRNRYCLARYEDLDVLAHINNTIN 1315
QY 1319 GGEKVIGVTRGAKSSLTGLPRINSAEGEIIIDGNIAKIGLHDLRFKITTIPDDPV 1378
DB 1319 GGEKVIGVTRGAKSSLTGLPRINSAEGEIIIDGNIAKIGLHDLRFKITTIPDDPV 1378
QY 1316 GGEKVIGVTRGAKSSLTGLPRINSAEGEIIIDGNIAKIGLHDLRFKITTIPDDPV 1375
DB 1316 GGEKVIGVTRGAKSSLTGLPRINSAEGEIIIDGNIAKIGLHDLRFKITTIPDDPV 1375
QY 1379 LFSGSLKMNLDPFQYDDEBWTSLLELAHLKDFVSALPDKLDHCEAGGENTLVSQORQV 1438
DB 1379 LFSGSLKMNLDPFQYDDEBWTSLLELAHLKDFVSALPDKLDHCEAGGENTLVSQORQV 1438
QY 1376 LFSGSLKMNLDPFQYDDEBWTSLLELAHLKDFVSALPDKLDHCEAGGENTLVSQORQV 1435
DB 1376 LFSGSLKMNLDPFQYDDEBWTSLLELAHLKDFVSALPDKLDHCEAGGENTLVSQORQV 1435
QY 1439 CLARALLRKTILVLDATAAVALDETDLIOSTIRTOFEDCTVLTIAHRLNTIMDYRVI 1498
DB 1439 CLARALLRKTILVLDATAAVALDETDLIOSTIRTOFEDCTVLTIAHRLNTIMDYRVI 1498
QY 1436 CLARALLRKTILVLDATAAVALDETDLIOSTIRTOFEDCTVLTIAHRLNTIMDYRVI 1495
DB 1436 CLARALLRKTILVLDATAAVALDETDLIOSTIRTOFEDCTVLTIAHRLNTIMDYRVI 1495
QY 1499 VLDKGEIOEYCAPSDLLQORGLFYSMAKDAGLV 1531
DB 1499 VLDKGEIOEYCAPSDLLQORGLFYSMAKDAGLV 1531
QY 1496 VLDKGEIOEYCAPSDLLQORGLFYSMAKDAGLV 1528
DB 1496 VLDKGEIOEYCAPSDLLQORGLFYSMAKDAGLV 1528

RESULT 15
US-08-462-109A-6
Sequence 6, Application US/08462109A
Patent No. 5882875
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,109A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Decont, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SRO ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-6

Query Match 70.7%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MARGFSSADGSDPLMNMNTNTNSNDPTKCFONTVTVVPCFYLMACPPFFLYLSRH 60
DB 1 MARSFSSADGSDPLMNMNTNTNSNDPTKCFONTVTVVPCFYLMACPPFFLYLSRH 60
QY 61 DRGYIOWTPLNKTITAGFLMIVCMADLFYSPMERSRGIFLAPVFLVSTLIGITTLA 120
DB 61 DRGYIOWTPLNKTITAGFLMIVCMADLFYSPMERSRGIFLAPVFLVSTLIGITTLA 120
QY 121 TFLIQERRKGVSSGIMLTFWLVALCALAILRSKIMTALKEBQVDFRDIITFYVFS 180
DB 121 TFLIQERRKGVSSGIMLTFWLVALCALAILRSKIMTALKEBQVDFRDIITFYVFS 180
QY 181 LLLIQVLSCFSRSPFLSETIHDNPNCPSSASFLSRITFWMITGLIYNGRQPLEGSD 240
DB 181 LLLIQVLSCFSRSPFLSETIHDNPNCPSSASFLSRITFWMITGLIYNGRQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKOPYKVYS-SKDPAPKSSKVDAEVEA 239
DB 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKOPYKVYS-SKDPAPKSSKVDAEVEA 239
QY 300 LIYKSPQKWNPSLFLVLYKTFPGPYFLMSFFKAIHDLMMFSGPOLIKLKEVNDTKAP 359
DB 300 LIYKSPQKWNPSLFLVLYKTFPGPYFLMSFFKAIHDLMMFSGPOLIKLKEVNDTKAP 359
QY 360 DMGGYRYTVLLPTTACIQITVHIOYRITCGVSGMRITKAVITAVYRKALVITNSARKST 419
DB 360 DMGGYRYTVLLPTTACIQITVHIOYRITCGVSGMRITKAVITAVYRKALVITNSARKST 419
QY 420 VGSIVNLMVDARFMDLATYINMISAPQVITIALTLMLNPGSVTLAGAVVMLVVP 479
DB 420 VGSIVNLMVDARFMDLATYINMISAPQVITIALTLMLNPGSVTLAGAVVMLVVP 479
QY 480 NAYMAKTKTYQVAHMKSKDNRIKLMNEIINGIKVLKLYVAMELAFQDKVLAIRQEBELKVL 539
DB 480 NAYMAKTKTYQVAHMKSKDNRIKLMNEIINGIKVLKLYVAMELAFQDKVLAIRQEBELKVL 539

QY 540 KKSAYLSAVGTFTWCTPPLVALCTPAVYVITDENNTLDAQTA FVSLAFNLIREFNITL 599
DB 541 KKSAYLSAVGTFTWCTPPLVALCTPAVYVITDENNTLDAQTA FVSLAFNLIREFNITL 600
QY 600 PMVSSIVQASVSLKRLRIFLSHELEPDSIERRPVQDGGTNSITVRNATFTARSDDP 659
DB 601 PMVSSIVQASVSLKRLRIFLSHELEPDSIERRSIKSGSG-NSITYKNTFTMARSEPP 659
QY 660 TLNGITFSIPEGALVAVVGVGCGKSSLSALIAEMDKVGHVAKGSVAVYPOQAMION 719
DB 660 TLNGITFSIPEGALVAVVGVGCGKSSLSALIAEMDKVGHVAKGSVAVYPOQAMION 719
QY 720 DSIRENTLFGHPIQENTYKAMBEACALLPLELILPSGDRTEIBKGYNLSGGQKORVSLA 779
DB 720 DSIRENTLFGHPIQENTYKAMBEACALLPLELILPSGDRTEIBKGYNLSGGQKORVSLA 779
QY 780 RAYVSNADYLPDDPLSAVDAAHGHKIFENVIGPKMKKKRTRILVTHSMGYLPQVDVYI 839
DB 780 RAYVSNADYLPDDPLSAVDAAHGHKIFENVIGPKMKKKRTRILVTHSMGYLPQVDVYI 839
QY 840 VMSGKISEMGSYOELLARDGAFAPFLRTYASTEOBODAEENGVTGSGKAKOMENG 899
DB 840 VMSGKISEMGSYOELLARDGAFAPFLRTYASTEOBODAEENGVTGSGKAKOMENG 899
QY 900 MLVTDAGKQLOPQLSSSSSSGDIRHNHNTAELOKAKKKEETWKLMAADAKOQGVK 959
DB 900 MLVTDAGKQLOPQLSSSSSSGDIRHNHNTAELOKAKKKEETWKLMAADAKOQGVK 959
QY 960 LSYVMDYMKRIGIFLSIFLPMCNHVSALASNYMLSLMTDD-PIYNGQOETHKRLSV 1018
DB 956 LSYVMDYMKRIGIFLSIFLPMCNHVSALASNYMLSLMTDD-PIYNGQOETHKRLSV 1015
QY 1019 YGALGISOGIAGVGYMAVSIIGIILASRCIHLVDLHLSILRSPMSFERTPSGNLVNRFSK 1078
DB 1016 YGALGISOGIAGVGYMAVSIIGIILASRCIHLVDLHLSILRSPMSFERTPSGNLVNRFSK 1075
QY 1079 ELDTYDSMTPEVIMKMGSLFNVTGACIYLATPIAIIIPPLGLIYFVQSPFYVASR 1138
DB 1076 ELDTYDSMTPEVIMKMGSLFNVTGACIYLATPIAIIIPPLGLIYFVQSPFYVASR 1135
QY 1139 QLRKLSVSRSPVYSHPNETLIGVSYIRAFEEQERFIHOSDLKVDENOKAYYPSIVANR 1198
DB 1136 QLRKLSVSRSPVYSHPNETLIGVSYIRAFEEQERFIHOSDLKVDENOKAYYPSIVANR 1195
QY 1199 LAVRELCVGNCTVLPALPAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNI 1258
DB 1196 LAVRELCVGNCTVLPALPAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNI 1255
QY 1259 VAVERLKEYSETKEAPMOIQETAPPSWPQGRVFEFRNYCLARYEDLDFVLRHIVTIN 1318
DB 1256 VAVERLKEYSETKEAPMOIQETAPPSWPQGRVFEFRNYCLARYEDLDFVLRHIVTIN 1315
QY 1319 GGEKVGIVGRTGAGKSLTLGLFRINESAGEIITIDGINIAKIGLHDLFKITIIPODPV 1378
DB 1316 GGEKVGIVGRTGAGKSLTLGLFRINESAGEIITIDGINIAKIGLHDLFKITIIPODPV 1375
QY 1379 LPSGSLRMLNDPESQYDEEVAWTSLELAHLKDFVSAIPDKLHBCAEGENLSVGOQOLV 1438
DB 1376 LPSGSLRMLNDPESQYDEEVAWTSLELAHLKDFVSAIPDKLHBCAEGENLSVGOQOLV 1435
QY 1439 CLARALLRKTIXLIVDEATAVLDLETDLIQSTIRPOFEDCYLTIARHLNTIMDYTRVI 1498
DB 1436 CLARALLRKTIXLIVDEATAVLDLETDLIQSTIRPOFEDCYLTIARHLNTIMDYTRVI 1495
QY 1499 VLDKGEIOEYGAAPSDLLQORGLFYSAKXDAGLV 1531
DB 1496 VLDKGEIOEYGAAPSDLLQORGLFYSAKXDAGLV 1528

Search completed: December 15, 2005, 15:49:32
Job time : 52.0811 secs

This page is blank

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:00 ; Search time 159,832 Seconds
(without alignments)
5037.528 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903

Sequence: 1 MALRGFCSADGSDPLMDMNV.....RSVAVAKKPKFSISPDLSL 1927

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9903	100.0	1927	US-10-665-283-1	Sequence 1, Appli
2	9891	99.9	1927	US-10-665-283-6	Sequence 6, Appli
3	9873	99.7	1947	US-10-665-283-8	Sequence 8, Appli
4	9734	98.3	1891	US-10-665-283-4	Sequence 4, Appli
5	7860	79.4	1531	US-09-939-853A-86	Sequence 86, Appli
6	7860	79.4	1531	US-10-667-891-6	Sequence 6, Appli
7	7860	79.4	1531	US-10-807-466-6	Sequence 6, Appli
8	7860	79.4	1531	US-10-484-577-678	Sequence 678, App
9	7860	79.4	1531	US-10-889-503-17	Sequence 19, Appli
10	7860	79.4	1531	US-10-756-149-5033	Sequence 5033, Ap
11	7769	78.5	1515	US-09-939-853A-87	Sequence 87, Appli
12	7763	78.1	1515	US-10-618-281-42	Sequence 42, Appli
13	7037	71.4	1388	US-10-408-765A-1718	Sequence 1718, Ap
14	7002.5	70.7	1528	US-09-939-853A-88	Sequence 88, Appli
15	6068.5	61.3	1303	US-10-618-281-63	Sequence 63, Appli
16	4866.5	49.1	1911	US-10-665-283-2	Sequence 2, Appli
17	4867.5	47.4	1875	US-10-665-283-5	Sequence 5, Appli
18	4667.5	47.1	1895	US-10-665-283-9	Sequence 9, Appli
19	4491.5	45.4	1901	US-10-665-283-10	Sequence 10, Appli
20	4491.5	45.4	1527	US-09-939-853A-83	Sequence 83, Appli
21	4491.5	45.4	1527	US-09-939-853A-84	Sequence 84, Appli
22	4491.5	45.4	1527	US-10-295-027-1342	Sequence 1342, Ap
23	4487.5	45.3	1530	US-10-889-503-6	Sequence 6, Appli
24	4487.5	45.3	1530	US-10-889-503-33	Sequence 33, Appli
25	4479.5	45.2	1530	US-10-889-503-55	Sequence 55, Appli
26	4396	44.4	1522	US-09-939-853A-85	Sequence 85, Appli
27	4135	41.8	1941	US-10-665-283-12	Sequence 12, Appli

28	4105	41.5	1961	5	US-10-665-283-7	Sequence 7, Appli
29	3966	40.0	1905	5	US-10-665-283-3	Sequence 3, Appli
30	3879.5	39.2	1548	5	US-10-807-466-28	Sequence 28, Appli
31	3860.5	39.0	1548	4	US-10-667-891-1	Sequence 1, Appli
32	3860.5	39.0	1548	5	US-10-807-466-1	Sequence 1, Appli
33	3846	38.8	1538	5	US-10-807-466-23	Sequence 23, Appli
34	3841	38.8	1538	5	US-10-807-466-24	Sequence 24, Appli
35	3835	38.7	1538	5	US-10-807-466-25	Sequence 25, Appli
36	3817.5	38.5	1539	5	US-10-807-466-26	Sequence 26, Appli
37	3817	38.5	1591	6	US-11-097-143-13266	Sequence 13266, A
38	3797	38.3	1538	5	US-10-807-466-27	Sequence 27, Appli
39	3779	38.2	1538	5	US-10-807-466-16	Sequence 16, Appli
40	3774	38.1	1538	5	US-10-807-466-18	Sequence 18, Appli
41	3773	38.1	1538	5	US-10-807-466-22	Sequence 22, Appli
42	3768	38.0	1538	5	US-10-807-466-17	Sequence 17, Appli
43	3754	37.9	1538	5	US-10-807-466-19	Sequence 19, Appli
44	3750.5	37.9	1539	5	US-10-807-466-20	Sequence 20, Appli
45	3733	37.7	1547	6	US-11-097-143-25248	Sequence 25248, A

ALIGNMENTS

RESULT 1									
US-10-665-283-1									
Sequence 1, Application US/10665283									
Publication No. US20050063989A1									
GENERAL INFORMATION:									
APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE									
DERAND, Renaud									
APPLICANT: GARCIA, Elisabeth									
PROST, Anne-Lise									
APPLICANT: REVILLAUD, Jean									
VIVAUDOU, Michel									
TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREO									
FILE REFERENCE: F263US104									
CURRENT APPLICATION NUMBER: US/10/665,283									
CURRENT FILING DATE: 2003-09-22									
NUMBER OF SEQ ID NOS: 25									
SOFTWARE: PatentIn version 3.1									
SEQ ID NO 1									
LENGTH: 1927									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-10-665-283-1									
Query Match									
Best Local Similarity 100.0%; Score 9903; DB 5; Length 1927;									
Matches 1927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MALRGFCSADGSDPLMDMNVNTNTSNDFTKCFQNTLVVWPCFYIMACFPFYLYSRH	60						
DB	1	MALRGFCSADGSDPLMDMNVNTNTSNDFTKCFQNTLVVWPCFYIMACFPFYLYSRH	60						
QY	61	DRGYIOWTPNKKTKTALGFLMTVCADLFFSFMERSRGIFLAPVPLVYSPTLIGITTLA	120						
DB	61	DRGYIOWTPNKKTKTALGFLMTVCADLFFSFMERSRGIFLAPVPLVYSPTLIGITTLA	120						
QY	121	TFPIQLERRRGVSGGIMLTFMVALVCALAIIRSKIMTALKEDAQVDLFRDITFYVYS	180						
DB	121	TFPIQLERRRGVSGGIMLTFMVALVCALAIIRSKIMTALKEDAQVDLFRDITFYVYS	180						
QY	181	LLLIQVLVSCFSDRSPLFSETIHDNPPCPSSASFLSRITFMWITGLIVRGQPLEGSD	240						
DB	181	LLLIQVLVSCFSDRSPLFSETIHDNPPCPSSASFLSRITFMWITGLIVRGQPLEGSD	240						
QY	241	LMSLNKEDTSEQVVPVIVKWKKECAKTRQPKVYVYSSKDPAPQPKSSSVVDANEVEAL	300						
DB	241	LMSLNKEDTSEQVVPVIVKWKKECAKTRQPKVYVYSSKDPAPQPKSSSVVDANEVEAL	300						
QY	301	IVKSPQKWNPSLFKVLKTFGPFYPLMSFPFKAIHDLMSFGQIILKLIKFVNDTRAPD	360						
DB	301	IVKSPQKWNPSLFKVLKTFGPFYPLMSFPFKAIHDLMSFGQIILKLIKFVNDTRAPD	360						

QY 361 WQGYFTVLLFVTAQTLVLHQQPHICFVSQMRKTAIVAGVYRKALVTINSARKSTV 420
DB 361 WQGYFTVLLFVTAQTLVLHQQPHICFVSQMRKTAIVAGVYRKALVTINSARKSTV 420
QY 421 GEIYNLMSVDAORFMDLATYTNMTWSAPLOVITLALYLMNLGSPVLAVGAVVMTLMPVN 480
DB 421 GEIYNLMSVDAORFMDLATYTNMTWSAPLOVITLALYLMNLGSPVLAVGAVVMTLMPVN 480
QY 481 AVAMAKTKTYOVANHKSKDNRIKLMNELINGIKVUKLYAMELAFKDKVLAIROBELKYLK 540
DB 481 AVAMAKTKTYOVANHKSKDNRIKLMNELINGIKVUKLYAMELAFKDKVLAIROBELKYLK 540
QY 541 KSAVLASAVGTFTWCTPPLVALCTPPAVYVTIDENNIIIDAQTA PVSALFNILRPLNITLP 600
DB 541 KSAVLASAVGTFTWCTPPLVALCTPPAVYVTIDENNIIIDAQTA PVSALFNILRPLNITLP 600
QY 601 MVISIVQASVSIKRIKRIEFLSHELEPDSIERPPKDGCTNSTTVRNATWTWASDDPT 660
DB 601 MVISIVQASVSIKRIKRIEFLSHELEPDSIERPPKDGCTNSTTVRNATWTWASDDPT 660
QY 661 LMGITFSPREGALVAVVGVQCGKSSLLSALLAEMDKEGHVAIKGSVAYVPOQAMIOND 720
DB 661 LMGITFSPREGALVAVVGVQCGKSSLLSALLAEMDKEGHVAIKGSVAYVPOQAMIOND 720
QY 721 SLRENILFEGCOLBEPYRSVIOACALLPDLLETPSGDRTEIGEGKVNLSGGQKQVSIAR 780
DB 721 SLRENILFEGCOLBEPYRSVIOACALLPDLLETPSGDRTEIGEGKVNLSGGQKQVSIAR 780
QY 781 AVYENADLYLEPDDPLSANDAHVGHKIFENVIGPKMKLNKRIILVTHMSNLYPOVDVITV 840
DB 781 AVYENADLYLEPDDPLSANDAHVGHKIFENVIGPKMKLNKRIILVTHMSNLYPOVDVITV 840
QY 841 MSGGKISEMGSYOELLARDGAFAPERTYASTEOBDAEENGCVTSVSGPKGAKOMENGM 900
DB 841 MSGGKISEMGSYOELLARDGAFAPERTYASTEOBDAEENGCVTSVSGPKGAKOMENGM 900
QY 901 LVTDGAKOQOROLSSSSSSSGDISRHNSSTAELQKAKKEETWKLMEADKAQTQVYL 960
DB 901 LVTDGAKOQOROLSSSSSSSGDISRHNSSTAELQKAKKEETWKLMEADKAQTQVYL 960
QY 961 SVYNDYMAKIGLIFISFLIFLPMCHVSAALSNYMLMTDTPRYNGOEHKRLSYYG 1020
DB 961 SVYNDYMAKIGLIFISFLIFLPMCHVSAALSNYMLMTDTPRYNGOEHKRLSYYG 1020
QY 1021 ALGISQGIAPFGYSAVASIGIILASRCJLVHDLHSILKSPMSFFERTPSGNLVNRFSEKL 1080
DB 1021 ALGISQGIAPFGYSAVASIGIILASRCJLVHDLHSILKSPMSFFERTPSGNLVNRFSEKL 1080
QY 1081 DTVMSEMIPEVIKMFNGSLFENVIGACIVILLATPIAIIIPPLGLYFFVQRFYVASSRQL 1140
DB 1081 DTVMSEMIPEVIKMFNGSLFENVIGACIVILLATPIAIIIPPLGLYFFVQRFYVASSRQL 1140
QY 1141 KRLESVSSPYSHNETLLGVSVTRAFEBQERFHOSDLKVDENQKAYPSIVANRWLA 1200
DB 1141 KRLESVSSPYSHNETLLGVSVTRAFEBQERFHOSDLKVDENQKAYPSIVANRWLA 1200
QY 1201 VRLBVCVACIVLFAALFAVISHSLSAGLVGSVSYSLQVTTYLWMLVMSSEMETNIVA 1260
DB 1201 VRLBVCVACIVLFAALFAVISHSLSAGLVGSVSYSLQVTTYLWMLVMSSEMETNIVA 1260
QY 1261 VERLKEVSETEKEADWQIOETAPSPSWPOVGRVEFRNYCLARYREDLFDVLARIHNTINGG 1320
DB 1261 VERLKEVSETEKEADWQIOETAPSPSWPOVGRVEFRNYCLARYREDLFDVLARIHNTINGG 1320
QY 1321 EKVGIVGRTGAGKSSLTGLFRINESABGEIIDGINIAKIGLHDLRFXITIIIPDDPVLF 1380
DB 1321 EKVGIVGRTGAGKSSLTGLFRINESABGEIIDGINIAKIGLHDLRFXITIIIPDDPVLF 1380
QY 1381 SGSLPMNLDPFSQYDEDEWVTSLELAHLKDPVSALPDKLDHBCAGGENTLVGQROVLCL 1440
DB 1381 SGSLPMNLDPFSQYDEDEWVTSLELAHLKDPVSALPDKLDHBCAGGENTLVGQROVLCL 1440

QY 1441 ARALIRKTKILVDEATAANDLEITDDLIQSTIRQEPDCTVLTARLNTIMDTYIVL 1500
DB 1441 ARALIRKTKILVDEATAANDLEITDDLIQSTIRQEPDCTVLTARLNTIMDTYIVL 1500
QY 1501 DKGEIOEYGA PSDLLOORGFLFYSNAKDAGLVGGGGGMLRKGIIPEEYVLTRLAEDPAE 1560
DB 1501 DKGEIOEYGA PSDLLOORGFLFYSNAKDAGLVGGGGGMLRKGIIPEEYVLTRLAEDPAE 1560
QY 1561 PRYRTRRRARFVSKKNCNVANHKNIREQGRFLQDVEFTTIVDLKMPHTLLIFTNSFLCSW 1620
DB 1561 PRYRTRRRARFVSKKNCNVANHKNIREQGRFLQDVEFTTIVDLKMPHTLLIFTNSFLCSW 1620
QY 1621 LLFPMWMLIAFAHGDLPAGEITVPCVTSIHSFSSAFLEISIEQVVTIGRGRRVTECP 1680
DB 1621 LLFPMWMLIAFAHGDLPAGEITVPCVTSIHSFSSAFLEISIEQVVTIGRGRRVTECP 1680
QY 1681 LAIILIVQNVGMINAVIMGCI FMTQAQHRAPETLIFSKAHVITLRHGRLCFMLRVG 1740
DB 1681 LAIILIVQNVGMINAVIMGCI FMTQAQHRAPETLIFSKAHVITLRHGRLCFMLRVG 1740
QY 1741 DLKRSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENGVGNGIFLVADLIYHYIDS 1800
DB 1741 DLKRSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENGVGNGIFLVADLIYHYIDS 1800
QY 1801 NSPIYDLAPSDLHHQDLEITVILBGEVETTGITTOARTSTYLADEILMGRFVPIVAED 1860
DB 1801 NSPIYDLAPSDLHHQDLEITVILBGEVETTGITTOARTSTYLADEILMGRFVPIVAED 1860
QY 1861 GRYSVDYSKFGNTIKVPTPLCTARQLEDERSILDALTLASSRGP LKRSVAVAKAPKFS 1920
DB 1861 GRYSVDYSKFGNTIKVPTPLCTARQLEDERSILDALTLASSRGP LKRSVAVAKAPKFS 1920

RESULT 2
US-10-665-283-6
; Sequence 6, Application US/10665283
; Publication No. US20050063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLIOD, Jean
; APPLICANT: VITAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: F263US104
; CURRENT APPLICATION NUMBER: US/10/665, 283
; CURRENT FILING DATE: 2003-09-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1927
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-6

Query Match 99.9%; Score 9891; DB 5; Length 1927;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCADSDPLMDNNTWNTNSDPTFCQONTVYVWPCFYIACPFPEFLYSRH 60
DB 1 MALRGFCADSDPLMDNNTWNTNSDPTFCQONTVYVWPCFYIACPFPEFLYSRH 60
QY 61 DRGYIOMPLNKTKTALGFLIMIVCMADLFYSFMSRSGIFLAPVFLVSPTLIGITTLA 120
DB 61 DRGYIOMPLNKTKTALGFLIMIVCMADLFYSFMSRSGIFLAPVFLVSPTLIGITTLA 120
QY 121 TFLIOLERRKGVQSSGIMLTFMLVALVCALALIRSKIMTALKEDAQVDLFRDITFFYVFS 180

Db 121 TFLQIERKGVQSSGIMLTFMVLVYCALAI LRSKINTALKEBAQVDFPDIIFYVFS 180
Qy 181 LLLIQVLVSCFSDSPFLSFETIHDNPNCPRESSASGLSTITWTMTGLIVRGYRQPLESSD 240
Db 181 LLLQVLVSCFSDSPFLSFETIHDNPNCPRESSASGLSTITWTMTGLIVRGYRQPLESSD 240
Qy 241 LMSLNKEDTSEQWVPLVYKMKKCKAKTRKQPVKVVYSSKDPAOPKSSKYDANEVVAL 300
Db 241 LMSLNKEDTSEQWVPLVYKMKKCKAKTRKQPVKVVYSSKDPAOPKSSKYDANEVVAL 300
Qy 301 IVKSPKEMNPSLFKVLKYTFGPYFLMSFPFKA IHDLMFSGPOLIKLII FVNDTKAPD 360
Db 301 IVKSPKEMNPSLFKVLKYTFGPYFLMSFPFKA IHDLMFSGPOLIKLII FVNDTKAPD 360
Qy 361 WQGFYFVTLFVYACLOTLVHOFPHICFVSGMRITKNAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYFVTLFVYACLOTLVHOFPHICFVSGMRITKNAVIGAVYRKALVITNSARKSSTV 420
Qy 421 GEIYNLMSVDAQRFMADLATYINMIMSAPIQVILALYILMLNGPSVLAVGAVMVLMPVN 480
Db 421 GEIYNLMSVDAQRFMADLATYINMIMSAPIQVILALYILMLNGPSVLAVGAVMVLMPVN 480
Qy 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDKVLAIROELKVLK 540
Db 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDKVLAIROELKVLK 540
Qy 541 KSAVLAVGTFVYCTPPLVALCTPNAVYITIDENNII DAOFAVSLAFENILRPPLNLP 600
Db 541 KSAVLAVGTFVYCTPPLVALCTPNAVYITIDENNII DAOFAVSLAFENILRPPLNLP 600
Qy 601 MVISITVOASVSLRLRIFLSHEELEPDSERRPVKGGGCTNSTITVNRATVTMARSDPT 660
Db 601 MVISITVOASVSLRLRIFLSHEELEPDSERRPVKGGGCTNSTITVNRATVTMARSDPT 660
Qy 661 LMGITFSP1PEGALVAVVGVGCGSKSLSLALAEWDKVEGVAIKGSVAYVPOQAMIOND 720
Db 661 LMGITFSP1PEGALVAVVGVGCGSKSLSLALAEWDKVEGVAIKGSVAYVPOQAMIOND 720
Qy 721 SLRENILFGCQLEPPYRSVIOACALLPDLIELPSGDRTEIGEKVNLSGGQKQVSLAR 780
Db 721 SLRENILFGCQLEPPYRSVIOACALLPDLIELPSGDRTEIGEKVNLSGGQKQVSLAR 780
Qy 781 AVYENADIYLFDDPLSAVDAVGHIFENVIIGPKMKLKNKRIILVYTHSMSTLPQVDYIIV 840
Db 781 AVYENADIYLFDDPLSAVDAVGHIFENVIIGPKMKLKNKRIILVYTHSMSTLPQVDYIIV 840
Qy 841 MSGGKISEMGSYOELIARDGAFAFLRTYASTBOEODAEENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYOELIARDGAFAFLRTYASTBOEODAEENGVTGVSQPGKEAKOMENGM 900
Qy 901 LVTDSACKOLQROLISSSSSYSGDISRRHNSTAELOKAEKKEETMKLEADKAQTGOVKL 960
Db 901 LVTDSACKOLQROLISSSSSYSGDISRRHNSTAELOKAEKKEETMKLEADKAQTGOVKL 960
Qy 961 SVYWDYKKAIGLFSFISIFLPMCNHVSALASNTWLSMTDDPIVNGTOEHTKTRLSVYG 1020
Db 961 SVYWDYKKAIGLFSFISIFLPMCNHVSALASNTWLSMTDDPIVNGTOEHTKTRLSVYG 1020
Qy 1021 ALGISOGIAVFGYMAVSIIGIILASRCIHDVLSIIRSPMSFERTPSGULVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGIILASRCIHDVLSIIRSPMSFERTPSGULVNRFSKEL 1080
Qy 1081 DTVDMSIPEVIKMFMSGLFNVIAGACIIVILATPIAIIIPGLIYFFVQRFYVASSROL 1140
Db 1081 DTVDMSIPEVIKMFMSGLFNVIAGACIIVILATPIAIIIPGLIYFFVQRFYVASSROL 1140
Qy 1141 KRLESVRSPPYSHFNETLLGVSVIRAFEEBERFIHOSDLKVDENOKAYVPSIYANRWLA 1200
Db 1141 KRLESVRSPPYSHFNETLLGVSVIRAFEEBERFIHOSDLKVDENOKAYVPSIYANRWLA 1200
Qy 1201 VRLECVCNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLVRMSSEMETNIVA 1260
Db 1201 VRLECVCNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLVRMSSEMETNIVA 1260

Db 1201 VRLECVCNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLVRMSSEMETNIVA 1260
Qy 1261 VERIKEYSETEKEMPMQIOETAPSSWPQVQVAFVFRNYCLARYREDLPVLBHINTVING 1320
Db 1261 VERIKEYSETEKEMPMQIOETAPSSWPQVQVAFVFRNYCLARYREDLPVLBHINTVING 1320
Qy 1321 EKVIGVBTGAGKSLTGLFRINESABGEI IIDGINAKIGLHDLRKITIIIPQDPVLF 1380
Db 1321 EKVIGVBTGAGKSLTGLFRINESABGEI IIDGINAKIGLHDLRKITIIIPQDPVLF 1380
Qy 1381 SGSLRNMULDEPSSQVSEDEWVTSLELAHLKDQVSAI PKLDHECAGEGEMLSVGQROLVCL 1440
Db 1381 SGSLRNMULDEPSSQVSEDEWVTSLELAHLKDQVSAI PKLDHECAGEGEMLSVGQROLVCL 1440
Qy 1441 ARALIRTKIIVLDEATAVADLETFDDLIIOSTIRPOFEDCTVLTIAHRLNTIMDYTRVIL 1500
Db 1441 ARALIRTKIIVLDEATAVADLETFDDLIIOSTIRPOFEDCTVLTIAHRLNTIMDYTRVIL 1500
Qy 1501 DKGEIOEYGA PSDLQORGLFYSAKXAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Db 1501 DKGEIOEYGA PSDLQORGLFYSAKXAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Qy 1561 PRYTRRRRARFVSKGNCVVAHNRNIREQGRFLQDVFTTVLDLKPHTLIIFTMSFLCSW 1620
Db 1561 PRYTRRRRARFVSKGNCVVAHNRNIREQGRFLQDVFTTVLDLKPHTLIIFTMSFLCSW 1620
Qy 1621 LLFPMVWMLIAFAHGLDAPGBGTNVCVTSIHSSSAPLSIEVQVITIGFGRNATECP 1680
Db 1621 LLFPMVWMLIAFAHGLDAPGBGTNVCVTSIHSSSAPLSIEVQVITIGFGRNATECP 1680
Qy 1681 LAILIILVONI VGINAINAMGCIEMKTAOAHRAEFLIISKAIVITLRHGRICFMRVYG 1740
Db 1681 LAILIILVONI VGINAINAMGCIEMKTAOAHRAEFLIISKAIVITLRHGRICFMRVYG 1740
Qy 1741 DLKSMIISATIIHQVVRKTTSPGEVAVPLHQVDIPNENGVGNGIFLVAPLIYHYIDS 1800
Db 1741 DLKSMIISATIIHQVVRKTTSPGEVAVPLHQVDIPNENGVGNGIFLVAPLIYHYIDS 1800
Qy 1801 NSPLYDLAPSDLHHHOLEIIVILEGEVETGITTQARTSYLADEIIMGQRFVIVABED 1860
Db 1801 NSPLYDLAPSDLHHHOLEIIVILEGEVETGITTQARTSYLADEIIMGQRFVIVABED 1860
Qy 1861 GRYSVDYSKRGNTIKVPTPLCTARQLEDDESLDALTLASRGPLRAASVAVAKKXKFS 1920
Db 1861 GRYSVDYSKRGNTIKVPTPLCTARQLEDDESLDALTLASRGPLRAASVAVAKKXKFS 1920
Qy 1921 ISPDLS 1927
Db 1921 ISPDLS 1927

RESULT 3
US-10-665-283-8
; Sequence 8, Application US/10665283
; Publication No. US20050063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLIOD, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF.
; FILE REFERENCE: P263US104
; CURRENT APPLICATION NUMBER: US/10/665,283
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-8

Query Match 99.7%; Score 9873; DB 5; Length 1947;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1927; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

```

Qy 1 MALRFGSADSDSLMDNMNTWNTNSPDFTKCFQNTVVMWPCFYMACEPFYELYSRH 60
Db 1 MALRFGSADSDSLMDNMNTWNTNSPDFTKCFQNTVVMWPCFYMACEPFYELYSRH 60
Qy 61 DRGIQMTPLNKKTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITTLA 120
Db 61 DRGIQMTPLNKKTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITTLA 120
Qy 61 DRGIQMTPLNKKTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITTLA 120
Db 61 DRGIQMTPLNKKTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITTLA 120
Qy 121 TFLIQLEBRKGVSSGIMLTFMLVALYCALAIRSKIMTALKEDAQVDLFPDITFYVYS 180
Db 121 TFLIQLEBRKGVSSGIMLTFMLVALYCALAIRSKIMTALKEDAQVDLFPDITFYVYS 180
Qy 181 LLLIQVLSCFSDSPFSETIHDNPPCBSSASFLSITFMWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLSCFSDSPFSETIHDNPPCBSSASFLSITFMWITGLIVRGYRQPLEGSD 240
Qy 241 LMSLNKEDTSBOVVPVLVKNMKCECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSBOVVPVLVKNMKCECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
Qy 301 IVKSPQKEMNPSLFKVLVYKTFGPYFLMSFFKRAIHDLMFSGPOILKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFKVLVYKTFGPYFLMSFFKRAIHDLMFSGPOILKLLIKFVNDTKAPD 360
Qy 361 WQGFYVTVLFTVACLQTLVHOYTHICFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGFYVTVLFTVACLQTLVHOYTHICFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDQRFMDLATYINMISAPLOYLIALYLMNLGSPVLAVMVLVNVN 480
Db 421 GEIYNLMSVDQRFMDLATYINMISAPLOYLIALYLMNLGSPVLAVMVLVNVN 480
Qy 481 AVMAKTKTYOVANHKSNDRIKLMNEILNGIKVLKLYAMELAFKQVLAIROBELKYLK 540
Db 481 AVMAKTKTYOVANHKSNDRIKLMNEILNGIKVLKLYAMELAFKQVLAIROBELKYLK 540
Qy 541 KSAVLASVGTFTWCTPFLVALCTPAVYVITDENNIDLAQTAFLVSLAFNLIRPLNLIP 600
Db 541 KSAVLASVGTFTWCTPFLVALCTPAVYVITDENNIDLAQTAFLVSLAFNLIRPLNLIP 600
Qy 601 MVISSIVQASVSLKRLRIFLSHEELPEPSIERRPVKDGGTNSITVRNATFTWASDPPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEELPEPSIERRPVKDGGTNSITVRNATFTWASDPPT 660
Qy 661 LMGITFSIPBGAVALVAVGVQVCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAMIOND 720
Db 661 LMGITFSIPBGAVALVAVGVQVCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAMIOND 720
Qy 721 SLRENIILFGCQLEBEYHRSVIOACALLPDLBILBEGDRTETGEKVMNSGGQKORVSLAR 780
Db 721 SLRENIILFGCQLEBEYHRSVIOACALLPDLBILBEGDRTETGEKVMNSGGQKORVSLAR 780
Qy 781 AVYSNADLYLPDDPLSAVDAAHVKHIFENVIKPKMKLNKTRILVTHSMVSLPOVDVITV 840
Db 781 AVYSNADLYLPDDPLSAVDAAHVKHIFENVIKPKMKLNKTRILVTHSMVSLPOVDVITV 840
Qy 841 MSGKISMGSVYOBELARDGAFAEFLRTYASTEQDABENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGKISMGSVYOBELARDGAFAEFLRTYASTEQDABENGVTGVSQPGKEAKOMENGM 900
Qy 901 LVYDSAGKOLOROUSSSSSYSGDISRHHNSTAELQKAPAKKEETWKLMEADKAOQGYKL 960
Db 901 LVYDSAGKOLOROUSSSSSYSGDISRHHNSTAELQKAPAKKEETWKLMEADKAOQGYKL 960
Qy 961 SVYDYMKAIGLFTSFLSIFLPMCHVASLASNYWLSLMTDPIVNGOETHKVLASYG 1020
Db 961 SVYDYMKAIGLFTSFLSIFLPMCHVASLASNYWLSLMTDPIVNGOETHKVLASYG 1020

```

```

Qy 1021 ALGISQIANFVGSYMANVSGIILASRCLHYDLHSILRSPMSFFERTPSGULVNRSEKEL 1080
Db 1021 ALGISQIANFVGSYMAVSGIILASRCLHYDLHSILRSPMSFFERTPSGULVNRSEKEL 1080
Qy 1081 DTVDSMIPEVIKPMGSLFVNVIGACIYILLATPIAIIIPPLGIYEFVORFYAASSROL 1140
Db 1081 DTVDSMIPEVIKPMGSLFVNVIGACIYILLATPIAIIIPPLGIYEFVORFYAASSROL 1140
Qy 1141 KRLSEVSRSPIVSHFNETLIGSVIRAFEBQEREIHOSDLKVDENQAKYPSIVANRWLA 1200
Db 1141 KRLSEVSRSPIVSHFNETLIGSVIRAFEBQEREIHOSDLKVDENQAKYPSIVANRWLA 1200
Qy 1201 VRLFCVNCIYLPALPAVISRHSLSAGLVLSYSLSQVTTYANMLVRMSSEMETIVA 1260
Db 1201 VRLFCVNCIYLPALPAVISRHSLSAGLVLSYSLSQVTTYANMLVRMSSEMETIVA 1260
Qy 1261 VERLKEYSETEKAPMIOETAPSSMPQVGRVFRNYCLRYREDLPVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPMIOETAPSSMPQVGRVFRNYCLRYREDLPVLRHINVTINGG 1320
Qy 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Qy 1381 SSGLRNMLDPPSOYSDEEWTSLFLAHLKDFVSLPKLDBHECABGEGNLSVGRQVCL 1440
Db 1381 SSGLRNMLDPPSOYSDEEWTSLFLAHLKDFVSLPKLDBHECABGEGNLSVGRQVCL 1440
Qy 1441 ABALRLKTKILVDEAFAVDELTDDLIOSTIRPQFEDCTVLTAAHLNLTMDTRYVL 1500
Db 1441 ABALRLKTKILVDEAFAVDELTDDLIOSTIRPQFEDCTVLTAAHLNLTMDTRYVL 1500
Qy 1501 DKGEIOEYGAPSDILOQRGLFYSSMAKADAGVGGGGGMLSRKGIIPREYVTLRLAEDPAE 1560
Db 1501 DKGEIOEYGAPSDILOQRGLFYSSMAKADAGVGGGGGMLSRKGIIPREYVTLRLAEDPAE 1560
Qy 1561 PRYTRRRRARFVSKKGCNVAKHNIQREOGRFLQDVFTVLVCLKMPTLLFTMSFLCSW 1620
Db 1561 PRYTRRRRARFVSKKGCNVAKHNIQREOGRFLQDVFTVLVCLKMPTLLFTMSFLCSW 1620
Qy 1621 LLFAMVWMLTAFAG-----DLAP-----GEGNVPCVTSHSSSAFLF 1680
Db 1621 LLFAMVWMLTAFAGDLIYAMEKGITDLAPPYVDPYAGGINVPCVTHSSSAFLF 1680
Qy 1681 SIEVOVITGFGGRNVTECPALILILIVONIVGMINAMLGCFMKTAAQARRAETLIF 1740
Db 1681 SIEVOVITGFGGRNVTECPALILILIVONIVGMINAMLGCFMKTAAQARRAETLIF 1740
Qy 1741 SKHAVITLHGRCLCFMLRVGDLRKSMTISATIHMQVVRKTTSPGGEVPLHQVDIPMENG 1800
Db 1741 SKHAVITLHGRCLCFMLRVGDLRKSMTISATIHMQVVRKTTSPGGEVPLHQVDIPMENG 1800
Qy 1801 VGGNGIFLAPLIIYHVIDNSPLYDLAPSDLHHODLEIIVILEGVEITGITTOARTS 1860
Db 1801 VGGNGIFLAPLIIYHVIDNSPLYDLAPSDLHHODLEIIVILEGVEITGITTOARTS 1860
Qy 1861 YLADEILMGQRFVIVAEEDRYSVDYSKFENITIKVPTLCTAQLDEDRSLDALTLAS 1920
Db 1861 YLADEILMGQRFVIVAEEDRYSVDYSKFENITIKVPTLCTAQLDEDRSLDALTLAS 1920

```

RESULT 4
 US-10-665-283-4
 ; Sequence 4, Application US/10665283
 ; Publication No. US2005063989A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
 ; APPLICANT: DERAND, Renaud
 ; APPLICANT: GARCIA, Elisabeth

APPLICANT: PROST, Anne-Lise
APPLICANT: REVILLOU, Jean
APPLICANT: VIVAUDOU, Michel
TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
FILE REFERENCE: F26305104
CURRENT APPLICATION NUMBER: US/10/665.283
NUMBER OF FILING DATE: 2003-09-22
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1891
TYPE: PRT
ORGANISM: Homo sapiens
US-10-665-283-4

Query Match 98.3%; Score 9734; DB 5; Length 1891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALGFGSADSDPLMDNMVMTNSNDPTFCQONTVYVWPCFYLACFPPEYLYLSRH 60
DB 1 MALGFGSADSDPLMDNMVMTNSNDPTFCQONTVYVWPCFYLACFPPEYLYLSRH 60

QY 61 DRGYIOMTPLNKTALGFLMIYCMADLFYSFMRSGIFLAPYFLVSPITLGIITLLA 120
DB 61 DRGYIOMTPLNKTALGFLMIYCMADLFYSFMRSGIFLAPYFLVSPITLGIITLLA 120

QY 121 TPLIQLERRKQVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLPFDITFYVFS 180
DB 121 TPLIQLERRKQVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLPFDITFYVFS 180

QY 181 LLLIOLVLSGDSPLSESTIHDPNCPBSSASLSTITFWMTGLIVRGYRQPLBESD 240
DB 181 LLLIOLVLSGDSPLSESTIHDPNCPBSSASLSTITFWMTGLIVRGYRQPLBESD 240

QY 241 LMSLNKEDTSEQVAVLVKMKKECAKTRKQPVKVVYSSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVAVLVKMKKECAKTRKQPVKVVYSSKDPAPKESKVDANEVEAL 300

QY 301 IVKSPQKEMNDSPFKVLYKTGPFYFLMSFFKAIHDLMMFSGPOLIKLIRFVNDTKAPD 360
DB 301 IVKSPQKEMNDSPFKVLYKTGPFYFLMSFFKAIHDLMMFSGPOLIKLIRFVNDTKAPD 360

QY 361 WQGFYVYVLLFVTAQLOTLVHOYFHLICFVSGMRKTAIVIGAVRKALVITNSARKSTV 420
DB 361 WQGFYVYVLLFVTAQLOTLVHOYFHLICFVSGMRKTAIVIGAVRKALVITNSARKSTV 420

QY 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNIGPSVLGAVVMTLMPVN 480
DB 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNIGPSVLGAVVMTLMPVN 480

QY 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLKYAMELAFKDKVLAIROBELKYLK 540
DB 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLKYAMELAFKDKVLAIROBELKYLK 540

QY 541 KSAIYLSAVGTFTWCTPPLVALCTPAVYVITIDENNILDAQAFVSLAFNLIRPLNLIP 600
DB 541 KSAIYLSAVGTFTWCTPPLVALCTPAVYVITIDENNILDAQAFVSLAFNLIRPLNLIP 600

QY 601 MVISSIVOASVSLKRLRIFLSHEELEPDSIERRPVKGCGGNSITVRNATVTMARSDPT 660
DB 601 MVISSIVOASVSLKRLRIFLSHEELEPDSIERRPVKGCGGNSITVRNATVTMARSDPT 660

QY 661 LMGITFSTIPEGALVAVVGQVCGKSSLSALLAEMDKVEGVAIKGSVAVYPOQAWIOND 720
DB 661 LMGITFSTIPEGALVAVVGQVCGKSSLSALLAEMDKVEGVAIKGSVAVYPOQAWIOND 720

QY 721 SLRNNIIFGCOLLEBPYRSVIOACALLPDLIELPSGDTEIGEGVNLISGQOKRVSLAR 780
DB 721 SLRNNIIFGCOLLEBPYRSVIOACALLPDLIELPSGDTEIGEGVNLISGQOKRVSLAR 780

QY 781 AVYSNADIYLPDDPLSAVDAHVKGHIFENVIGPKMKLNKTRILVTHSMSTYLPQVDVILV 840
DB 781 AVYSNADIYLPDDPLSAVDAHVKGHIFENVIGPKMKLNKTRILVTHSMSTYLPQVDVILV 840

DB 781 AVYSNADIYLPDDPLSAVDAHVKGHIFENVIGPKMKLNKTRILVTHSMSTYLPQVDVILV 840

QY 841 MSGGKISMSGYOBLLARDGAFAPFLRYATSTBEOAEENGVTGSPRGAEAKOMENG 900
DB 841 MSGGKISMSGYOBLLARDGAFAPFLRYATSTBEOAEENGVTGSPRGAEAKOMENG 900

QY 901 LVTSAGKOLOROSSSSSSSGDISRHNSPAELOKAEKAEETWKMEADAKAQTQYKL 960
DB 901 LVTSAGKOLOROSSSSSSSGDISRHNSPAELOKAEKAEETWKMEADAKAQTQYKL 960

QY 961 SVYWDYKAIQGLFISPLSIFLMCNHVSALASNWLSLMTDPIVNGTOBHTKRLSYVG 1020
DB 961 SVYWDYKAIQGLFISPLSIFLMCNHVSALASNWLSLMTDPIVNGTOBHTKRLSYVG 1020

QY 1021 ALGISOGIAYGVGMAVSIIGITLASRCLHVDLHSILRSPMSFFERTPSGULVNRFSKEL 1080
DB 1021 ALGISOGIAYGVGMAVSIIGITLASRCLHVDLHSILRSPMSFFERTPSGULVNRFSKEL 1080

QY 1081 DTVDSMIPEVTKMFMGSLFENVIGACIVILATPIAIIIPPLGLIYFFVORFYVASSRQI 1140
DB 1081 DTVDSMIPEVTKMFMGSLFENVIGACIVILATPIAIIIPPLGLIYFFVORFYVASSRQI 1140

QY 1141 KRLESVSRSPYSHFNETLGVSVIRAFEEQERFIHOSDLKYDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNETLGVSVIRAFEEQERFIHOSDLKYDENOKAYPSIVANRWLA 1200

QY 1201 VRLCEVNCIYLPALPAVTSRHSLSAGLVLSYSLSQTTTYNMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALPAVTSRHSLSAGLVLSYSLSQTTTYNMLVRMSSEMETNIVA 1260

QY 1261 VERLKEYSETEKAPMOIOETAPSSWPQVRVFRNYCLARYEDLPVLRPHIVNTNGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSWPQVRVFRNYCLARYEDLPVLRPHIVNTNGG 1320

QY 1321 EKVGIIVGRTGAGKSLITGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIIVGRTGAGKSLITGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380

QY 1381 SGSLRNMULDPSQYSDSEWVTSLELAHLKDFVSALPKLDHECAEGGENLSVGRQIYCL 1440
DB 1381 SGSLRNMULDPSQYSDSEWVTSLELAHLKDFVSALPKLDHECAEGGENLSVGRQIYCL 1440

QY 1441 ARALLRKTIVLDEATAVDLETPDLIOSTRIRQFEDCTVLTAAHLNITMDTRYIVL 1500
DB 1441 ARALLRKTIVLDEATAVDLETPDLIOSTRIRQFEDCTVLTAAHLNITMDTRYIVL 1500

QY 1501 DKGEIOBYGAPSDLLQORGLFYSMAXDAGLVGGGGGMLSRKGIIPBEYVUTRIABDPAE 1560
DB 1501 DKGEIOBYGAPSDLLQORGLFYSMAXDAGLVGGGGGMLSRKGIIPBEYVUTRIABDPAE 1560

QY 1561 PRYTRERRARFVSKKNCNVAAHKNIREQGRFIQDVFVTLVDLKWPHULLIFTNMPLCSW 1620
DB 1561 PRYTRERRARFVSKKNCNVAAHKNIREQGRFIQDVFVTLVDLKWPHULLIFTNMPLCSW 1620

QY 1621 LIPFAMVWMLIAFAAGDLAPBGTNPVCTVSHSPSSAFPLSIEQVATIGFGRMVTBECR 1680
DB 1621 LIPFAMVWMLIAFAAGDLAPBGTNPVCTVSHSPSSAFPLSIEQVATIGFGRMVTBECR 1680

QY 1681 LAIILILVONIVGMINAIVMGCIIFMKTAAHRAETLIFSKAIVITLRHGRCLCFMLRVG 1740
DB 1681 LAIILILVONIVGMINAIVMGCIIFMKTAAHRAETLIFSKAIVITLRHGRCLCFMLRVG 1740

QY 1741 DLRSKMIISATIHMOVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAPLIYHVIDS 1800
DB 1741 DLRSKMIISATIHMOVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAPLIYHVIDS 1800

QY 1801 NSPLYDLAPSDLHHHOLEIIVILEGVENTGTITTOARTSYLDELIMGQRFVPTVAEED 1860
DB 1801 NSPLYDLAPSDLHHHOLEIIVILEGVENTGTITTOARTSYLDELIMGQRFVPTVAEED 1860

QY 1861 GRYSVDYSKFGNTIKVPTPLCTARQLDEDRS 1891
DB 1861 GRYSVDYSKFGNTIKVPTPLCTARQLDEDRS 1891

RESULT 5
US-09-939-853A-86
; Sequence 86, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20/27,337
; PRIOR APPLICATION NUMBER: 60/277,337
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-86

Query Match 79.4%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALNGFCASDSDPLMDNMVMTNTNSPDTKCFONTVLVWVCFYLMACFPEFYLYLSRH 60
DB 1 MALNGFCASDSDPLMDNMVMTNTNSPDTKCFONTVLVWVCFYLMACFPEFYLYLSRH 60
QY 61 DRGIQYMTPLNKTATGELLMTVCADLFYSFWEBSRGIFLAPVFLVSPITLGIITLLA 120
DB 61 DRGIQYMTPLNKTATGELLMTVCADLFYSFWEBSRGIFLAPVFLVSPITLGIITLLA 120
QY 61 DRGIQYMTPLNKTATGELLMTVCADLFYSFWEBSRGIFLAPVFLVSPITLGIITLLA 120
DB 61 DRGIQYMTPLNKTATGELLMTVCADLFYSFWEBSRGIFLAPVFLVSPITLGIITLLA 120
QY 121 TFLQLERKRGVQSSGIMLTFWLVALVCAALIRKIMTALKEQOVDFRDTTYVVS 180
DB 121 TFLQLERKRGVQSSGIMLTFWLVALVCAALIRKIMTALKEQOVDFRDTTYVVS 180
QY 121 TFLQLERKRGVQSSGIMLTFWLVALVCAALIRKIMTALKEQOVDFRDTTYVVS 180
DB 121 TFLQLERKRGVQSSGIMLTFWLVALVCAALIRKIMTALKEQOVDFRDTTYVVS 180
QY 181 LLLIQLVSCFSDSPLFSETHDNPCESSASFLSRITTFWIMGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVSCFSDSPLFSETHDNPCESSASFLSRITTFWIMGLIVRGYRQPLEGSD 240
QY 181 LLLIQLVSCFSDSPLFSETHDNPCESSASFLSRITTFWIMGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVSCFSDSPLFSETHDNPCESSASFLSRITTFWIMGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPQKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPQKSSKVDANEVEAL 300
QY 301 IVKSPQKEMNPSLFVLYKTGPFLMSPFPAIHDLMMFGSPQILKLIKRVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFVLYKTGPFLMSPFPAIHDLMMFGSPQILKLIKRVNDTKAPD 360
QY 301 IVKSPQKEMNPSLFVLYKTGPFLMSPFPAIHDLMMFGSPQILKLIKRVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFVLYKTGPFLMSPFPAIHDLMMFGSPQILKLIKRVNDTKAPD 360
QY 361 WQGYFYTYLLFVTAACLOTLYLHQYFHCIVSGMRKTKTAVIGAVYKALVITNSARKSSTV 420
DB 361 WQGYFYTYLLFVTAACLOTLYLHQYFHCIVSGMRKTKTAVIGAVYKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLLMLNGLPSVLGAVAVMLVAVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLLMLNGLPSVLGAVAVMLVAVN 480
QY 481 AVMAKTKTYOYAHMKSKONRIKLMNEILNGIKVLTAYAMELAFQKYLAIQOELKYLK 540
DB 481 AVMAKTKTYOYAHMKSKONRIKLMNEILNGIKVLTAYAMELAFQKYLAIQOELKYLK 540
QY 541 KSAVLSAAGTFTWCTPFLVALCFEAVYVTDENNIIILAAQTAFFVSLALFNILRPPLNLP 600
DB 541 KSAVLSAAGTFTWCTPFLVALCFEAVYVTDENNIIILAAQTAFFVSLALFNILRPPLNLP 600
QY 601 MVISIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDDEPT 660

DB 601 MVISIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDDEPT 660
QY 661 LMGITFSEIPEGALVAVVGOVCGKSSLSLALAEADKVEGVAIKGSVAVYPQOAWIOND 720
DB 661 LMGITFSEIPEGALVAVVGOVCGKSSLSLALAEADKVEGVAIKGSVAVYPQOAWIOND 720
QY 721 SLRENILFGCQLEPPYRSVIOACALLPDLIELPSGRTEIGEKGVLNLSGGOKORVSLAR 780
DB 721 SLRENILFGCQLEPPYRSVIOACALLPDLIELPSGRTEIGEKGVLNLSGGOKORVSLAR 780
QY 781 AVYNNADIIYLEDPLSAVDAAVGHIFENYVGPCKMLKNKRIILVYMSVLPQVDVIV 840
DB 781 AVYNNADIIYLEDPLSAVDAAVGHIFENYVGPCKMLKNKRIILVYMSVLPQVDVIV 840
QY 841 MSGGKISEMGSYQELLARDGAFAELRTYASTQEQODAEENGVTGVSGBPGEAKOMENGM 900
DB 841 MSGGKISEMGSYQELLARDGAFAELRTYASTQEQODAEENGVTGVSGBPGEAKOMENGM 900
QY 901 LVYDSAGKOLQROLSSSSSYSGDISRHNSYAEIQAAPAKKEFTWKLMEADKAQTGVYKL 960
DB 901 LVYDSAGKOLQROLSSSSSYSGDISRHNSYAEIQAAPAKKEFTWKLMEADKAQTGVYKL 960
QY 961 SVYWDYMKATGLFISFLSIFLPMCNHVSALASNYWLSLMTDDPIVNGTQOHTKVLSSYG 1020
DB 961 SVYWDYMKATGLFISFLSIFLPMCNHVSALASNYWLSLMTDDPIVNGTQOHTKVLSSYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTSGMLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTSGMLVNRFSKEL 1080
QY 1081 DTVDMSIPEVIKMGSLFNIVIGACIVILLATPIAIIIPPLGIYFFVQGFYVASSRQL 1140
DB 1081 DTVDMSIPEVIKMGSLFNIVIGACIVILLATPIAIIIPPLGIYFFVQGFYVASSRQL 1140
QY 1141 KRLSSVSSPYSHFNETLLGVSVIRAFEEQEREIHOSDLKVDNOKAYPSIVANRWLA 1200
DB 1141 KRLSSVSSPYSHFNETLLGVSVIRAFEEQEREIHOSDLKVDNOKAYPSIVANRWLA 1200
QY 1201 VRLCEVNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVFRNYCLARYRDLDFLHINVTINGG 1320
DB 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVFRNYCLARYRDLDFLHINVTINGG 1320
QY 1321 EKVGIYGRGTGAKSSLTGLFRINESAGEIITIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIYGRGTGAKSSLTGLFRINESAGEIITIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGLRMLNDDPFSQVSDDEWMTSLSLAHLKDFVSLPDKLDHECAEGENLSVGOBOLVCL 1440
DB 1381 SGLRMLNDDPFSQVSDDEWMTSLSLAHLKDFVSLPDKLDHECAEGENLSVGOBOLVCL 1440
QY 1441 ARAILRKTKIIVLDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ARAILRKTKIIVLDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOEYGAQSDLLQORGLFYSMADAGLV 1531
DB 1501 DKGEIOEYGAQSDLLQORGLFYSMADAGLV 1531

RESULT 6
US-10-667-891-6
; Sequence 6, Application US/10667891
; Publication No. US20040171024A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BREY, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRAILLES, MARINE

APPLICANT: RZHEITSKY, ANDREY
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
FILE REFERENCE: 03495-0294-00000
CURRENT APPLICATION NUMBER: US/10/667,891
CURRENT FILING DATE: 2003-09-23
PRIOR APPLICATION NUMBER: 60/413,469
PRIOR FILING DATE: 2002-09-26
NUMBER OF SEQ ID NOS: 76
SOFTWARE: Patentin Ver. 3.2
SEQ ID NO: 6
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo sapiens
US-10-667-891-6

Query Match 79.4%; Score 7860; DB 4; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFSADGSDPLMDMNTMTNTPFTKCFQNTYLVWPCYTLNACPFYFLYSRH 60
DB 1 MALRGFSADGSDPLMDMNTMTNTPFTKCFQNTYLVWPCYTLNACPFYFLYSRH 60
QY 61 DRGYIOMTPLNKTALGFLMIYCMADLFYSFMEGRSGIFLAVFVLSPTLIGITLLA 120
DB 61 DRGYIOMTPLNKTALGFLMIYCMADLFYSFMEGRSGIFLAVFVLSPTLIGITLLA 120
QY 121 TFLIOLERRKGVSGSGLMFLVALVCAALILRSKIMTALKEDAQVDFRDIFFYYFS 180
DB 121 TFLIOLERRKGVSGSGLMFLVALVCAALILRSKIMTALKEDAQVDFRDIFFYYFS 180
QY 181 LLLIOLVLSGSDSPFSETIHDNPNCPSSSASFSLRTITWTGILVGRYOLPESD 240
DB 181 LLLIOLVLSGSDSPFSETIHDNPNCPSSSASFSLRTITWTGILVGRYOLPESD 240
QY 181 LLLIOLVLSGSDSPFSETIHDNPNCPSSSASFSLRTITWTGILVGRYOLPESD 240
DB 181 LLLIOLVLSGSDSPFSETIHDNPNCPSSSASFSLRTITWTGILVGRYOLPESD 240
QY 241 LMSLNKEDTSEOVVYVVKWKKECACTRKOPVYVYSSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSEOVVYVVKWKKECACTRKOPVYVYSSKDPAPKESKVDANEVEAL 300
QY 301 IVKSPQKEMNPSEFKVLKTFGPYFLMSFEFKAIHDLMSGPOILKLLIKFVNDTRAPD 360
DB 301 IVKSPQKEMNPSEFKVLKTFGPYFLMSFEFKAIHDLMSGPOILKLLIKFVNDTRAPD 360
QY 361 WOGFYTVLLFVTRACLOTLYLHQYFHCIFVSGMRKTAIVGAYVRKALVITNSARKSSTV 420
DB 361 WOGFYTVLLFVTRACLOTLYLHQYFHCIFVSGMRKTAIVGAYVRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAORPMDLATTYINMTWSAPLOYILALYLLMLTGPSTVAGVAVMLAMPVN 480
DB 421 GEIYNLMSVDAORPMDLATTYINMTWSAPLOYILALYLLMLTGPSTVAGVAVMLAMPVN 480
QY 481 AVNMMKTITQVAMHMSKDNRIKLMNEILNGIKVLYLAMELAFKDKVLAIRBELKVLK 540
DB 481 AVNMMKTITQVAMHMSKDNRIKLMNEILNGIKVLYLAMELAFKDKVLAIRBELKVLK 540
QY 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITDENNNIIDAQTAFAVSLAFNITLREPLNLP 600
DB 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITDENNNIIDAQTAFAVSLAFNITLREPLNLP 600
QY 601 MVSISIVQASVSLKRLIFLSHELEPDSIERPVKGGGNSITVNNATFTWARSPT 660
DB 601 MVSISIVQASVSLKRLIFLSHELEPDSIERPVKGGGNSITVNNATFTWARSPT 660
QY 661 LINGITFIPGALVAVVGCGCKSSLLSALLAEMDVESGVAIKGSVAVVPOQAWIOND 720
DB 661 LINGITFIPGALVAVVGCGCKSSLLSALLAEMDVESGVAIKGSVAVVPOQAWIOND 720
QY 721 SLRENIIFGCOLLEPPYRSYIQAICALPDLEILPSGRTEIGEKNVLSGGQKORVSLAR 780
DB 721 SLRENIIFGCOLLEPPYRSYIQAICALPDLEILPSGRTEIGEKNVLSGGQKORVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDHVGKIFENVIGPKMLKNKTRILVTHSMYSYLPQVDYIIV 840

DB 781 AVYSNADIYLFDDPLSAVDHVGKIFENVIGPKMLKNKTRILVTHSMYSYLPQVDYIIV 840
QY 841 MSGGKISEMGSYOBELARDCAFELRTYASTOBDAENGTYGSGPKEKAKOMENG 900
DB 841 MSGGKISEMGSYOBELARDCAFELRTYASTOBDAENGTYGSGPKEKAKOMENG 900
QY 901 LVTDSAGKOLQORLSSSSSGDISRHNSHSTAELQRAEAKKEFTWKLMEADKQOTGVKL 960
DB 901 LVTDSAGKOLQORLSSSSSGDISRHNSHSTAELQRAEAKKEFTWKLMEADKQOTGVKL 960
QY 961 SYVMDYKKAIGLFISFLIFMCHNYSALASNYMLSLMTDDEIYVNGTOHTKRLSVYG 1020
DB 961 SYVMDYKKAIGLFISFLIFMCHNYSALASNYMLSLMTDDEIYVNGTOHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSGIILASRCLHVDLHSLIRSPMSFPERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGYSMAVSGIILASRCLHVDLHSLIRSPMSFPERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMIPVYIKMFMGSLFVIGACTVILATPILATIPPLGLIYFVQRFYVASSROL 1140
DB 1081 DTVDSMIPVYIKMFMGSLFVIGACTVILATPILATIPPLGLIYFVQRFYVASSROL 1140
QY 1141 KRLSEYSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSEYSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCTIVLPAALFAVISRSLSAGLVGSVSLQYTYTLNMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVGNCTIVLPAALFAVISRSLSAGLVGSVSLQYTYTLNMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSMPQVGRFERNYCLRYEEDLDFVLRHINVTING 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSMPQVGRFERNYCLRYEEDLDFVLRHINVTING 1320
QY 1321 EKVIGVGRTAGSSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVIGVGRTAGSSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
QY 1381 SGLRNLNLDPSQYSDSEEVTSLELAHLKDFVSALPDKLDHECAGEGENISVGOQLVCL 1440
DB 1381 SGLRNLNLDPSQYSDSEEVTSLELAHLKDFVSALPDKLDHECAGEGENISVGOQLVCL 1440
QY 1441 ARALLKTKTLVNDERTAAVDETDLIOSTITTOEDCTVLTARLNTIMTYRIVL 1500
DB 1441 ARALLKTKTLVNDERTAAVDETDLIOSTITTOEDCTVLTARLNTIMTYRIVL 1500
QY 1501 DKGEIOEYGAPSDLOORGIFYMAKDAQIV 1531
DB 1501 DKGEIOEYGAPSDLOORGIFYMAKDAQIV 1531

RESULT 7
US-10-807-466-6
Sequence 6, Application US/10807466
Publication No. US20040244066A1
GENERAL INFORMATION:
APPLICANT: ROTH, CHARLES W.
APPLICANT: BREY, PAUL T.
APPLICANT: HOLM, INGE
APPLICANT: GRALLIES, MARINE
APPLICANT: RZHEITSKY, ANDREY
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
FILE REFERENCE: 03495-0294-01000
CURRENT APPLICATION NUMBER: US/10/807,466
CURRENT FILING DATE: 2004-03-24
PRIOR APPLICATION NUMBER: 10/667,891
PRIOR FILING DATE: 2003-09-23
PRIOR APPLICATION NUMBER: 60/413,469
PRIOR FILING DATE: 2002-09-26
NUMBER OF SEQ ID NOS: 139
SOFTWARE: Patentin Ver. 3.2

```
; SEQ ID NO 6
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-807-466-6

Query Match      79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALRGFCSDGSDPLMDMNTMTNSNPDFTKCFONTVLVWVPCFYLMACFPFFFLYLSRH 60
Db      1 MALRGFCSDGSDPLMDMNTMTNSNPDFTKCFONTVLVWVPCFYLMACFPFFFLYLSRH 60
Qy      61 DRGIQWTPLNKKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGTTLLA 120
Db      61 DRGIQWTPLNKKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGTTLLA 120
Qy      121 TFLIQLEERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLPFDITFYVYS 180
Db      121 TFLIQLEERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLPFDITFYVYS 180
Qy      181 LLLIQVLVSCFSDSPSEFETHDNPCESSASFLSITFMWTGLIVRGYRQPLBSSD 240
Db      181 LLLIQVLVSCFSDSPSEFETHDNPCESSASFLSITFMWTGLIVRGYRQPLBSSD 240
Qy      241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVVAL 300
Db      241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVVAL 300
Qy      301 IVKSPQKEMNPDLFKVLYKTGPYFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPD 360
Db      301 IVKSPQKEMNPDLFKVLYKTGPYFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPD 360
Qy      361 MGGFFYTLTLVVTCLQTLVHOVHIFGVSGMRKKTAVIGAVYKALVITNSARKSTV 420
Db      361 MGGFFYTLTLVVTCLQTLVHOVHIFGVSGMRKKTAVIGAVYKALVITNSARKSTV 420
Qy      421 GEIYVNLVSVDARFMDLATYINMIWSAPLOVILALYLLMLNGSPVLGAVAVMVLMEVN 480
Db      421 GEIYVNLVSVDARFMDLATYINMIWSAPLOVILALYLLMLNGSPVLGAVAVMVLMEVN 480
Qy      481 AVMAKTKTYOVAAHKSKONRIKLMMELINGIKVLKLYAMELAFKQVLAIRQBELKVLK 540
Db      481 AVMAKTKTYOVAAHKSKONRIKLMMELINGIKVLKLYAMELAFKQVLAIRQBELKVLK 540
Qy      541 KSAVLSAAGTFTWCTPFLVALCTPAVYVTIDENNILDAQTAFFSLFNLIRFPLNLLP 600
Db      541 KSAVLSAAGTFTWCTPFLVALCTPAVYVTIDENNILDAQTAFFSLFNLIRFPLNLLP 600
Qy      601 MVISSIVQASVSLKRLIFLSHEBLEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
Db      601 MVISSIVQASVSLKRLIFLSHEBLEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
Qy      661 LVGTFSPISPGALVAVVGQVGGKSSLLSALLAEMDKVEGHVAKGSVAAYVPQQAIOND 720
Db      661 LVGTFSPISPGALVAVVGQVGGKSSLLSALLAEMDKVEGHVAKGSVAAYVPQQAIOND 720
Qy      721 SLRENIILGCGQLEERYRSVTOACALLPDLBITLPSGDRTEIGEKVNSGGCKORVSLAR 780
Db      721 SLRENIILGCGQLEERYRSVTOACALLPDLBITLPSGDRTEIGEKVNSGGCKORVSLAR 780
Qy      781 AVYSNADLYLPDDPLSAVDAAHVGKHIIFENVIGPKGMLKNKTRILVTHMSYLPQVDVIV 840
Db      781 AVYSNADLYLPDDPLSAVDAAHVGKHIIFENVIGPKGMLKNKTRILVTHMSYLPQVDVIV 840
Qy      841 MSGKISMGSYOELLARDGAFAEFLRTYASTBOBDAEENGVTGVSQPKAKOMENGM 900
Db      841 MSGKISMGSYOELLARDGAFAEFLRTYASTBOBDAEENGVTGVSQPKAKOMENGM 900
Qy      901 LVTSBAGQOLQOLSSSSSSYSQDISRHNSTALEOKAEKKKEETKMLMEADQAQGOYKL 960
Db      901 LVTSBAGQOLQOLSSSSSSYSQDISRHNSTALEOKAEKKKEETKMLMEADQAQGOYKL 960
```

```
Qy      961 SVYWDYKKAIGLFLSFLSIFLPMCNHVSALASNYWLSLMTDPIVNGTOEHTKRLSVYG 1020
Db      961 SVYWDYKKAIGLFLSFLSIFLPMCNHVSALASNYWLSLMTDPIVNGTOEHTKRLSVYG 1020
Qy      1021 ALGISOGIAVFGYMAVSTIGIILASRCLHYDLHSILRSWSPFEPTPSGTLVNRFSKEL 1080
Db      1021 ALGISOGIAVFGYMAVSTIGIILASRCLHYDLHSILRSWSPFEPTPSGTLVNRFSKEL 1080
Qy      1081 DTVDSMIPEVYKMGMSLFPNVIGACIVILATPIAIIIPPLGIFYFVQRFYASSROL 1140
Db      1081 DTVDSMIPEVYKMGMSLFPNVIGACIVILATPIAIIIPPLGIFYFVQRFYASSROL 1140
Qy      1141 KRLSVSRSPYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQAYYPSIVANFWLA 1200
Db      1141 KRLSVSRSPYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQAYYPSIVANFWLA 1200
Qy      1201 VRLCEVNCIVLPALFAVVISRHSLSAGVLGVSVYSLOVTTYLNMVLRMSSEMETIVA 1260
Db      1201 VRLCEVNCIVLPALFAVVISRHSLSAGVLGVSVYSLOVTTYLNMVLRMSSEMETIVA 1260
Qy      1261 VERLKEYSETEKEAPMOIOETAPPSWPOVGRVFRNYCLYREDLPFLRHINVTINGG 1320
Db      1261 VERLKEYSETEKEAPMOIOETAPPSWPOVGRVFRNYCLYREDLPFLRHINVTINGG 1320
Qy      1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKTTIIPQDPVLF 1380
Db      1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKTTIIPQDPVLF 1380
Qy      1381 SGIIRMLDPPSQSDEEWTSLTSLAHUKDFVSLPKLHCEKGEGBENLSVGOROLVCL 1440
Db      1381 SGIIRMLDPPSQSDEEWTSLTSLAHUKDFVSLPKLHCEKGEGBENLSVGOROLVCL 1440
Qy      1441 ARALLRKTKILVDEATAVVDLEETDILQSTIRFQFEDCTVLTAHRLNTIMDYTRVIL 1500
Db      1441 ARALLRKTKILVDEATAVVDLEETDILQSTIRFQFEDCTVLTAHRLNTIMDYTRVIL 1500
Qy      1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
Db      1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 8
US-10-484-577-678
; Sequence 678, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft
; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGT1A
; FILE REFERENCE: P2285PCT-1
; CURRENT APPLICATION NUMBER: US/10/484,577
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: PCT/EP 02/08220
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: EP 01 11 7608.8
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: EP 02011710.7
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 683
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 678
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-484-577-678

Query Match      79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALRGFCSDGSDPLMDMNTMTNSNPDFTKCFONTVLVWVPCFYLMACFPFFFLYLSRH 60
Db      1 MALRGFCSDGSDPLMDMNTMTNSNPDFTKCFONTVLVWVPCFYLMACFPFFFLYLSRH 60
```

QY 61 DRGYIOMTPLNKTALGFLMIVCMADLFYSFWERSRGIPLAPVLYSPULLGITTLLA 120
DB 61 DRGYIOMTPLNKTALGFLMIVCMADLFYSFWERSRGIPLAPVLYSPULLGITTLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMVALVICALILRSKIMTALKEDAOVDLFRDITFYVYFS 180
DB 121 TFLIOLERRKGVSSGIMLTFMVALVICALILRSKIMTALKEDAOVDLFRDITFYVYFS 180
QY 181 LLLIQLVLSGSDSPFLSETIHDNPNCPRESSASFLSITITWTGLIVRGYROPLEGSD 240
DB 181 LLLIQLVLSGSDSPFLSETIHDNPNCPRESSASFLSITITWTGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSBOVVPVLYKMKKECAKTRKOPVAVVYSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSBOVVPVLYKMKKECAKTRKOPVAVVYSSKDPAPKSSKVDANEVEAL 300
QY 301 IVKSPKEMWNSLKFVLYKTFPGPYFLMSFPFKALHDLMPFGPOLKLLIKFVNDTKAPD 360
DB 301 IVKSPKEMWNSLKFVLYKTFPGPYFLMSFPFKALHDLMPFGPOLKLLIKFVNDTKAPD 360
QY 361 WOGFYTVLLFVTKCLQTLVLHQYFHCFSGMRKITAIVIGAVYRKALVITNSARKSSTV 420
DB 361 WOGFYTVLLFVTKCLQTLVLHQYFHCFSGMRKITAIVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAORFMDLATYINMTWSAPIOYIILALYILMTLIGSVYLAGVAVMLAMPVN 480
DB 421 GEIYNLMSVDAORFMDLATYINMTWSAPIOYIILALYILMTLIGSVYLAGVAVMLAMPVN 480
QY 481 AVNAMKTXYOVAMHMSKDNRIKLMNEILINGIKLYAMELAFKDKVLAIRBELKYLK 540
DB 481 AVNAMKTXYOVAMHMSKDNRIKLMNEILINGIKLYAMELAFKDKVLAIRBELKYLK 540
QY 541 KSAVLSAVGFTWVCTPFLVALCTFAVYVITDENNII DAOFAVLSALFNILRPPLNLP 600
DB 541 KSAVLSAVGFTWVCTPFLVALCTFAVYVITDENNII DAOFAVLSALFNILRPPLNLP 600
QY 601 MVISSIVQASVSLKRLIFLSHELEPDSIERPVKCGGNGSTVNNAFTTAKRSPT 660
DB 601 MVISSIVQASVSLKRLIFLSHELEPDSIERPVKCGGNGSTVNNAFTTAKRSPT 660
QY 661 LINGITFSIPGALVAVVGVCGGKSLSLALAEMLKVEGHVAKGSVAVYPOQAMQND 720
DB 661 LINGITFSIPGALVAVVGVCGGKSLSLALAEMLKVEGHVAKGSVAVYPOQAMQND 720
QY 721 SLRENIIFGQLEBPYRSYIOACALLPDLIELPSGRTIEGKGVNLSSGQKQVSLAR 780
DB 721 SLRENIIFGQLEBPYRSYIOACALLPDLIELPSGRTIEGKGVNLSSGQKQVSLAR 780
QY 781 AVVSNMDIYLFDDPLSAVDAHNGKHI FENVIGPKGMKAKTRILVTHSMGYLPOVDYIIV 840
DB 781 AVVSNMDIYLFDDPLSAVDAHNGKHI FENVIGPKGMKAKTRILVTHSMGYLPOVDYIIV 840
QY 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOBDAENGVTGVSQPKGAQOMENGM 900
DB 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOBDAENGVTGVSQPKGAQOMENGM 900
QY 901 LVTDSAGKQLOROQLSSSSSYSGDISRHNSYTAELQKAKEKEFTWKLMEADKAQTQGVKL 960
DB 901 LVTDSAGKQLOROQLSSSSSYSGDISRHNSYTAELQKAKEKEFTWKLMEADKAQTQGVKL 960
QY 961 SVYMDWYKATGLFISPLSTFLFMCNHYSAASNYWLSLWDDPVLNCTOHTKVRISLVG 1020
DB 961 SVYMDWYKATGLFISPLSTFLFMCNHYSAASNYWLSLWDDPVLNCTOHTKVRISLVG 1020
QY 1021 ALGISOGIAVFGYSMAVISGIIILASRCIAYDLHLSIRSPMSFPERTPSGNLVNRSEKEL 1080
DB 1021 ALGISOGIAVFGYSMAVISGIIILASRCIAYDLHLSIRSPMSFPERTPSGNLVNRSEKEL 1080
QY 1081 DTYDSMIPYIYKFMGSLENVIGACIYIILATPIAIIIPPLGIYFFVORFYVASSROL 1140
DB 1081 DTYDSMIPYIYKFMGSLENVIGACIYIILATPIAIIIPPLGIYFFVORFYVASSROL 1140
QY 1140 DTYDSMIPYIYKFMGSLENVIGACIYIILATPIAIIIPPLGIYFFVORFYVASSROL 1140
DB 1140 DTYDSMIPYIYKFMGSLENVIGACIYIILATPIAIIIPPLGIYFFVORFYVASSROL 1140

QY 1141 KRLSEVSRPYSHFNETLLGVSVYIRAFEBEORERTHOSDLKVDENOKAYVPSIVANRWLA 1200
DB 1141 KRLSEVSRPYSHFNETLLGVSVYIRAFEBEORERTHOSDLKVDENOKAYVPSIVANRWLA 1200
QY 1201 VRLCEVNCIYLFALFAVIRSRHLSAGLVGLSVYSLOVTTYLNMVVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLFALFAVIRSRHLSAGLVGLSVYSLOVTTYLNMVVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPPSWPQYRVEFRNYCLRYREDDLPVLRIHVVTTNGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPPSWPQYRVEFRNYCLRYREDDLPVLRIHVVTTNGG 1320
QY 1321 EKVIGVGTAGKSSSLTGLFRINESAGEIILIDGINIAKIGHDLAFKTIIPQDPVLF 1380
DB 1321 EKVIGVGTAGKSSSLTGLFRINESAGEIILIDGINIAKIGHDLAFKTIIPQDPVLF 1380
QY 1381 SGSLRNMNLDPPSOYSDSEWVTSLELAHLKDQFVSLPDKLHCEAGGENTSVGQROLVCL 1440
DB 1381 SGSLRNMNLDPPSOYSDSEWVTSLELAHLKDQFVSLPDKLHCEAGGENTSVGQROLVCL 1440
QY 1441 APALLRKTKLIVDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
DB 1441 APALLRKTKLIVDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
QY 1501 DKGEIOEYGAAPSDLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAAPSDLQORGLFYSMAKDAGLV 1531

RESULT 9
US-10-889-503-19
; Sequence 19, Application US/10889503
; Publication No. US20050063968A1
; GENERAL INFORMATION:
; APPLICANT: Fox Chase Cancer Center
; APPLICANT: Kruth, Gary D.
; APPLICANT: Lee, Kun
; APPLICANT: Belinsky, Martin G.
; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
; FILE REFERENCE: FCCC 98-02
; CURRENT APPLICATION NUMBER: US/10/889,503
; PRIOR FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: US/09/647,140
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US99/06644
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/079,759
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/095,153
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-889-503-19
Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCASDGSPLMDMNTVMTSNPDFTKCFQONTVLVWPCFYLMACPPFYLYLSRH 60
DB 1 MALRGFCASDGSPLMDMNTVMTSNPDFTKCFQONTVLVWPCFYLMACPPFYLYLSRH 60
QY 61 DRGYIOMTPLNKTALGFLMIVCMADLFYSFWERSRGIPLAPVLYSPULLGITTLLA 120
DB 61 DRGYIOMTPLNKTALGFLMIVCMADLFYSFWERSRGIPLAPVLYSPULLGITTLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMVALVICALILRSKIMTALKEDAOVDLFRDITFYVYFS 180
DB 121 TFLIOLERRKGVSSGIMLTFMVALVICALILRSKIMTALKEDAOVDLFRDITFYVYFS 180

121 TFLQLERRKGVSSGIMLTFMWALVALCALALRSKIMTALKEADAQVDLFFDITFYVFS 180
181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSRITFMWITGLIVRGYRQPLEGSD 240
181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSRITFMWITGLIVRGYRQPLEGSD 240
241 LMSINKEDTSEQVVPVIVKMKKECAKTRKQPVKVVVSSKDPAPKSSKVDANEVEAL 300
241 LMSINKEDTSEQVVPVIVKMKKECAKTRKQPVKVVVSSKDPAPKSSKVDANEVEAL 300
301 IVSPKEMNPSLKFVLYKTGPFYFLMSFFPKAIDHLMFSGPOLKLLIFVNDTYAPD 360
301 IVSPKEMNPSLKFVLYKTGPFYFLMSFFPKAIDHLMFSGPOLKLLIFVNDTYAPD 360
361 WOGFYTVLFLVTAQCLQTLVLAHQFHLICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420
361 WOGFYTVLFLVTAQCLQTLVLAHQFHLICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420
421 GEIVNLMSVDAQRFMDLATYINMIMSAPLOVILALYLLMLNGPSVLGAVAVMLMVEVN 480
421 GEIVNLMSVDAQRFMDLATYINMIMSAPLOVILALYLLMLNGPSVLGAVAVMLMVEVN 480
481 AVAMANKTKYOVAMKSKDNRIKLMNETLNGITKYLKYAMELAPKDYLAIRQBELKYLK 540
481 AVAMANKTKYOVAMKSKDNRIKLMNETLNGITKYLKYAMELAPKDYLAIRQBELKYLK 540
541 KSAVLASVGTFTWCTPPLVALCTFAVAVYTIDENNIDDAQTAFAVSLAFNIIIRPLNLP 600
541 KSAVLASVGTFTWCTPPLVALCTFAVAVYTIDENNIDDAQTAFAVSLAFNIIIRPLNLP 600
601 MVISSIVQASVSLKRLRIFLSHEELPEPDSERRPVKDGGSITVRNATFTWARSDEPT 660
601 MVISSIVQASVSLKRLRIFLSHEELPEPDSERRPVKDGGSITVRNATFTWARSDEPT 660
661 LMGITFTSPREGALVAVVGQVCGGSSLLSALLAENDKXEGHVAIKGSAAVYPOQAMIOND 720
661 LMGITFTSPREGALVAVVGQVCGGSSLLSALLAENDKXEGHVAIKGSAAVYPOQAMIOND 720
721 SLRNNILFGCOLPEPYRSVIOACALPDLLELPSGDTEIGEKGVLNSGGOKORVSLAR 780
721 SLRNNILFGCOLPEPYRSVIOACALPDLLELPSGDTEIGEKGVLNSGGOKORVSLAR 780
781 AVYNSADIYLFDDPLSAVDAHVGHIFENVIQPKGMLNKKTRILVTHSMSTYLPQVDVIV 840
781 AVYNSADIYLFDDPLSAVDAHVGHIFENVIQPKGMLNKKTRILVTHSMSTYLPQVDVIV 840
841 MSGKISMSGYSQELIARDGAFAEFLRTYASTEQDAEENGVTGVSQPGKEAKOMENGM 900
841 MSGKISMSGYSQELIARDGAFAEFLRTYASTEQDAEENGVTGVSQPGKEAKOMENGM 900
901 LVTOAGKOLOROSSSSSSSGDISRHNSSTAELOKAKKEBTMKLEADKAOTGOVKL 960
901 LVTOAGKOLOROSSSSSSSGDISRHNSSTAELOKAKKEBTMKLEADKAOTGOVKL 960
961 SVYNDYMKAIQLFISFLSIFLPMCNSVALASNTWLSMTDPIVNGTOEHTKVLSTYYG 1020
961 SVYNDYMKAIQLFISFLSIFLPMCNSVALASNTWLSMTDPIVNGTOEHTKVLSTYYG 1020
1021 ALGISQGIATVFGYMAVSIIGIILASRCLHVDLHSLASPMSPFRRTPSGULVNFSEKL 1080
1021 ALGISQGIATVFGYMAVSIIGIILASRCLHVDLHSLASPMSPFRRTPSGULVNFSEKL 1080
1081 DTVOSMIDPEVIMKMFGLFENVIGACIVLLATPIAIIIPPLGLIYFFVQRYVASSKQL 1140
1081 DTVOSMIDPEVIMKMFGLFENVIGACIVLLATPIAIIIPPLGLIYFFVQRYVASSKQL 1140
1141 KRLESVSRSPYVSHFNFTLGVSVIRAREBOERFIHQSDLKVDENQKAYYSIVANRMLA 1200
1141 KRLESVSRSPYVSHFNFTLGVSVIRAREBOERFIHQSDLKVDENQKAYYSIVANRMLA 1200
1201 VRLECVGNCIVLPALFAVISRHSLSAGLVGLSVSYLQVTTYLNMLVRMSSEMETNIVA 1260
1201 VRLECVGNCIVLPALFAVISRHSLSAGLVGLSVSYLQVTTYLNMLVRMSSEMETNIVA 1260

1201 VRLECVGNCIVLPALFAVISRHSLSAGLVGLSVSYLQVTTYLNMLVRMSSEMETNIVA 1260
1261 VERLKESETEKEAPWQIOETAPSSWPQYGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320
1261 VERLKESETEKEAPWQIOETAPSSWPQYGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320
1321 EKVGIVERTGAGSSLTGLFRINESAGEIIGDINIAICGLHDLARFKITITIPQDPVLF 1380
1321 EKVGIVERTGAGSSLTGLFRINESAGEIIGDINIAICGLHDLARFKITITIPQDPVLF 1380
1381 SSGLRMNLDPFSQVSDSEWNTSLELAHLKDVFSAALPKLDHECAGEGENISVGOROLVCL 1440
1381 SSGLRMNLDPFSQVSDSEWNTSLELAHLKDVFSAALPKLDHECAGEGENISVGOROLVCL 1440
1441 ABALARKTKILVDEATAVADLETDLIQSTIRPOEDCTVLATIANRIANTIMDYTRYVL 1500
1441 ABALARKTKILVDEATAVADLETDLIQSTIRPOEDCTVLATIANRIANTIMDYTRYVL 1500
1501 DKGEOYGAUSDLLQORGLFYSMADAGLV 1531
1501 DKGEOYGAUSDLLQORGLFYSMADAGLV 1531
RESULT 10
US-10-756-149-5033
; Sequence 5033, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Azitok, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: Patent version 3.2
; SEQ ID NO 5033
; LENGTH: 1531
; TYPE: PRN
; ORGANISM: Homo Sapiens
US-10-756-149-5033
Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MALRGFCSADGSDPLMDMNTWTNTSNDFPKCFQNTVLAWVPCRYLMACPFFFLYLSRH 60
1 MALRGFCSADGSDPLMDMNTWTNTSNDFPKCFQNTVLAWVPCRYLMACPFFFLYLSRH 60
61 DRGYIOMTPLNKTATLGFLLMIVCMADLFYSFWEBSRGIFLAVPLVSPTLGITTLA 120
61 DRGYIOMTPLNKTATLGFLLMIVCMADLFYSFWEBSRGIFLAVPLVSPTLGITTLA 120
61 DRGYIOMTPLNKTATLGFLLMIVCMADLFYSFWEBSRGIFLAVPLVSPTLGITTLA 120
121 TFLQLERRKGVSSGIMLTFMWALVALCALALRSKIMTALKEADAQVDLFFDITFYVFS 180
121 TFLQLERRKGVSSGIMLTFMWALVALCALALRSKIMTALKEADAQVDLFFDITFYVFS 180
181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSRITFMWITGLIVRGYRQPLEGSD 240
181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSRITFMWITGLIVRGYRQPLEGSD 240
181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSRITFMWITGLIVRGYRQPLEGSD 240
241 LMSINKEDTSEQVVPVIVKMKKECAKTRKQPVKVVVSSKDPAPKSSKVDANEVEAL 300
241 LMSINKEDTSEQVVPVIVKMKKECAKTRKQPVKVVVSSKDPAPKSSKVDANEVEAL 300
301 IVSPKEMNPSLKFVLYKTGPFYFLMSFFPKAIDHLMFSGPOLKLLIFVNDTYAPD 360
301 IVSPKEMNPSLKFVLYKTGPFYFLMSFFPKAIDHLMFSGPOLKLLIFVNDTYAPD 360
361 WOGFYTVLFLVTAQCLQTLVLAHQFHLICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420
361 WOGFYTVLFLVTAQCLQTLVLAHQFHLICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420

Db 361 WQGYFTVLLEVTACLOTVLHQQYFHLCEVSGMRKTAIVGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLMSVPAORFMDLATYINMTMSA.PIQVTLALYLLMTLNGSVLAGVAVMTLMPVN 480
Db 421 GEIVNLMSVPAORFMDLATYINMTMSA.PIQVTLALYLLMTLNGSVLAGVAVMTLMPVN 480
Qy 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIRQBELKVLK 540
Db 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIRQBELKVLK 540
Qy 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITDENNNILDAQAFVSLAFNLIRFPLNLP 600
Db 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITDENNNILDAQAFVSLAFNLIRFPLNLP 600
Qy 601 MVSSIVQASVSLKRLAIFLSHELEPDSIERPVYKGGGNSITVNNATPTAKRSPPPT 660
Db 601 MVSSIVQASVSLKRLAIFLSHELEPDSIERPVYKGGGNSITVNNATPTAKRSPPPT 660
Qy 661 LNCITFSGIPREGALVAVVGVCGKSSLSLALLAEMDKVEGHVATKGSVAVYPOQAMIOND 720
Db 661 LNCITFSGIPREGALVAVVGVCGKSSLSLALLAEMDKVEGHVATKGSVAVYPOQAMIOND 720
Qy 721 SLRENILFGCQLEBPYRSVIAQCALLPDLIELPSGDRTEIGEKVNLSGGQKQVSLAR 780
Db 721 SLRENILFGCQLEBPYRSVIAQCALLPDLIELPSGDRTEIGEKVNLSGGQKQVSLAR 780
Qy 781 AVYSNADIIYLFDDPLSAVDHVGKHTPENYIGPGMKLKNKRIILVTHSMSTLPQVDYIIV 840
Db 781 AVYSNADIIYLFDDPLSAVDHVGKHTPENYIGPGMKLKNKRIILVTHSMSTLPQVDYIIV 840
Qy 841 MSGGKISEMGSYOELRLARDGAFELRTYASTEOEDAEENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYOELRLARDGAFELRTYASTEOEDAEENGVTGVSQPGKEAKOMENGM 900
Qy 901 LVTDLSAGKQLOQROSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKQOTQOVKL 960
Db 901 LVTDLSAGKQLOQROSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKQOTQOVKL 960
Qy 961 SVYWDYKAKIQLFISFLIFPMCNHVSALASNYWLSLMTDDPIVNGTOHTKRLSVYG 1020
Db 961 SVYWDYKAKIQLFISFLIFPMCNHVSALASNYWLSLMTDDPIVNGTOHTKRLSVYG 1020
Qy 1021 ALGISOGIAVFGVMAVSIIGIILASRCILHVDLHSILSRPMSPEERPSGNLVNRFSEKL 1080
Db 1021 ALGISOGIAVFGVMAVSIIGIILASRCILHVDLHSILSRPMSPEERPSGNLVNRFSEKL 1080
Qy 1081 DTVDSMIPVYIKPMGSLFNVIGACIYIILATPIAIIIPPLGLIYFVQRFVYASSROL 1140
Db 1081 DTVDSMIPVYIKPMGSLFNVIGACIYIILATPIAIIIPPLGLIYFVQRFVYASSROL 1140
Qy 1141 KRLBSVSRSPVYSHFNFTLLGVSVYIRAFEEQERFIHQSDLKVDENQKAYYPSIVANTWLA 1200
Db 1141 KRLBSVSRSPVYSHFNFTLLGVSVYIRAFEEQERFIHQSDLKVDENQKAYYPSIVANTWLA 1200
Qy 1201 VRLBECVNCIYLPALFAVISRHSLSAGLVGLSVSYLQVTTYLNMVLRMSSEMETNIVA 1260
Db 1201 VRLBECVNCIYLPALFAVISRHSLSAGLVGLSVSYLQVTTYLNMVLRMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKAPMOIQETAPPSWPQVGRVFRNYCJRYREDDLFVLRHINVTNGG 1320
Db 1261 VERLKEVSETEKAPMOIQETAPPSWPQVGRVFRNYCJRYREDDLFVLRHINVTNGG 1320
Qy 1321 EKVGIVERTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
Db 1321 EKVGIVERTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
Qy 1381 SSGSRNMULDPSOYSDSEWVTSLEIHLKDFVSLPKLHDECAEGEENISVQGRQVLC 1440
Db 1381 SSGSRNMULDPSOYSDSEWVTSLEIHLKDFVSLPKLHDECAEGEENISVQGRQVLC 1440
Qy 1441 ARALLRKTKILVDEATAVADLETDDLIQSTIRTOFEDCTVLTAHLANTIMDYTRYIVL 1500
Db 1441 ARALLRKTKILVDEATAVADLETDDLIQSTIRTOFEDCTVLTAHLANTIMDYTRYIVL 1500

Qy 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531

RESULT 11
US-09-939-853A-87
; Sequence 87, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT FILING DATE: US/09/939, 853A
; PRIOR APPLICATION NUMBER: 2001-08-27
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-87

Query Match 78.5%; Score 7769; DB 3; Length 1515;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 DNVVNTSNPDPTKCRQNVLVWVPCFYLMACPPFYLYLSRHHGCIQMTPLNKTCTA 76
Db 1 DNVVNTSNPDPTKCRQNVLVWVPCFYLMACPPFYLYLSRHHGCIQMTPLNKTCTA 60
Qy 77 LGFLMIVCADDIFYSEWERSRGIPLAPVPLVSPTLIGITTLATFLIQLERRKGVQSSG 136
Db 61 LGFLMIVCADDIFYSEWERSRGIPLAPVPLVSPTLIGITTLATFLIQLERRKGVQSSG 120
Qy 137 IMTFMLVALVCAILIRSKIMTALXEDAVDLPDITFYVYFSLIIQVLISCFSDRSP 196
Db 121 IMTFMLVALVCAILIRSKIMTALXEDAVDLPDITFYVYFSLIIQVLISCFSDRSP 180
Qy 197 LFSETHDPRCPRESSASFLSRITTFWITGLYRGYRQPLEGSDMSLNKEDTSEOVVPV 256
Db 181 LFSETHDPRCPRESSASFLSRITTFWITGLYRGYRQPLEGSDMSLNKEDTSEOVVPV 240
Qy 257 LVNWMKKECAKTKOPKVVYVSSKDPAPQKSSKVDANEVEALIVKSPKEMNPISLFKY 316
Db 241 LVNWMKKECAKTKOPKVVYVSSKDPAPQKSSKVDANEVEALIVKSPKEMNPISLFKY 300
Qy 317 LYTFGPYFLMSEFPFRAIHDLMFSGPOIILKLIKFTVNDTKADWQGYFTVLLPYTACL 376
Db 301 LYTFGPYFLMSEFPFRAIHDLMFSGPOIILKLIKFTVNDTKADWQGYFTVLLPYTACL 360
Qy 377 QTVLHQQYFHCIVSGKRIKTAIVGAVYRKALVITNSARKSSVGEIVNLMSVPAORFMD 436
Db 361 QTVLHQQYFHCIVSGKRIKTAIVGAVYRKALVITNSARKSSVGEIVNLMSVPAORFMD 420
Qy 437 LATYINMISAPQVITLALYLLNLGSPVLAGVAVVNLVAVPNAVMAKTKTYQVAAHMK 496
Db 421 LATYINMISAPQVITLALYLLNLGSPVLAGVAVVNLVAVPNAVMAKTKTYQVAAHMK 480
Qy 497 SKONRIKLMNEILNGIKVLKLYAMELAFKDKVLAIRQBELKVLKSAVAVGTFTWCTV 556
Db 481 SKONRIKLMNEILNGIKVLKLYAMELAFKDKVLAIRQBELKVLKSAVAVGTFTWCTV 540
Qy 557 PFLVALCTFAVYVITDENNNILDAQFVSLAFNLIRFPLNLPMTISSIVQASVSLKRL 616

|||||
Db 541 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSALFNILRPPLNILPMVSISSIVQASVSKRL 600
Qy 617 RIFLSHELEPDSIERRPVKDGGGNSITVNNATFTWARSPTLNGITFSIPGALVAV 676
Db 601 RIFLSHELEPDSIERRPVKDGGGNSITVNNATFTWARSPTLNGITFSIPGALVAV 660
Qy 677 VGVGCGKSSLSLALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGCQLEBPY 736
Db 661 VGVGCGKSSLSLALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGCQLEBPY 720
Qy 737 YRSYIQAACALLPDLIELIPSGDRTEIGEKNVLSGGQKORVSLARAAYSNADIYFDDPLS 796
Db 721 YRSYIQAACALLPDLIELIPSGDRTEIGEKNVLSGGQKORVSLARAAYSNADIYFDDPLS 780
Qy 797 AYDAHVGHKHFENYIGPKMKLNKTRILVTHSMSTLPVDVITVMSGGKISEMSYOELL 856
Db 781 AYDAHVGHKHFENYIGPKMKLNKTRILVTHSMSTLPVDVITVMSGGKISEMSYOELL 840
Qy 857 ARDAFAEFLLRTYASTEOEDAENGVTGVSQPKAKOMENGLVTDASAGKOLQROUSS 916
Db 841 ARDAFAEFLLRTYASTEOEDAENGVTGVSQPKAKOMENGLVTDASAGKOLQROUSS 900
Qy 917 SSSYSGDISRHHNSTAEIOKAKEETWKLMEADKAQTGOVKLSVYWDYMKALGLFTSF 976
Db 901 SSSYSGDISRHHNSTAEIOKAKEETWKLMEADKAQTGOVKLSVYWDYMKALGLFTSF 960
Qy 977 LSTFLPMONHVSALASNYWLSLMTDDPIVNGTOEHKTRLSVYGALGISQGIAYGVGMA 1036
Db 961 LSTFLPMONHVSALASNYWLSLMTDDPIVNGTOEHKTRLSVYGALGISQGIAYGVGMA 1020
Qy 1037 VSIIGIILASRCLAHYDLHLSIRSPMSPEERTPSGNLVNRFSEKEDTVDMSIPEVIMKFMG 1096
Db 1021 VSIIGIILASRCLAHYDLHLSIRSPMSPEERTPSGNLVNRFSEKEDTVDMSIPEVIMKFMG 1080
Qy 1097 SLFNVIAGCIYIILATPIAIIIPPLGLIYFFVQRFYVASSRQKRLSVSRSPYSHFN 1156
Db 1081 SLFNVIAGCIYIILATPIAIIIPPLGLIYFFVQRFYVASSRQKRLSVSRSPYSHFN 1140
Qy 1157 ETLIGVSVIRAFEEGERFIHOSDLKVDENOKAYPSIVANRWMLAVRLCVCNCTVLPAL 1216
Db 1141 ETLIGVSVIRAFEEGERFIHOSDLKVDENOKAYPSIVANRWMLAVRLCVCNCTVLPAL 1200
Qy 1217 FAVISRHSLSAGVLGVLSYSLOVTTYLNMLVRMSSEMETNI VAVERLKEYSETEKEAPW 1276
Db 1201 FAVISRHSLSAGVLGVLSYSLOVTTYLNMLVRMSSEMETNI VAVERLKEYSETEKEAPW 1260
Qy 1277 QIOETAPSSWPQVGRVFEFRNYCLARYREDLDFVLHINVTINGEKVGI VERTGAGKSSL 1336
Db 1261 QIOETAPSSWPQVGRVFEFRNYCLARYREDLDFVLHINVTINGEKVGI VERTGAGKSSL 1320
Qy 1337 TLGFRINESAGEIIIDIGINIAKGLHDLRPKTIITIODPVLFGSGSLRMLNDPSSQYSD 1396
Db 1321 TLGFRINESAGEIIIDIGINIAKGLHDLRPKTIITIODPVLFGSGSLRMLNDPSSQYSD 1380
Qy 1397 EEWVTSLELAHKDVFVSLPKLDHECABGGENLSVGOQOLVCLARALLRKTKILVLEBA 1456
Db 1381 EEWVTSLELAHKDVFVSLPKLDHECABGGENLSVGOQOLVCLARALLRKTKILVLEBA 1440
Qy 1457 TAAVLETFDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAESDILQ 1516
Db 1441 TAAVLETFDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAESDILQ 1500
Qy 1517 ORGLFYSMAXDAGLV 1531
Db 1501 ORGLFYSMAXDAGLV 1515

RESULT 12

US-10-618-281-42

; Sequence 42, Application US/10618281

; Publication No. US20040219609A1

; GENERAL INFORMATION:

; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618, 281
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1515
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-618-281-42

Query Match 78.4%; Score 7763; DB 5; Length 1515;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 DNANTNTSNPDFTKCQNTYLVWVPCFYLMACPPFFLYLSRDRGYIQMTPLNKTKTA 76
Db 1 DNANTNTSNPDFTKCQNTYLVWVPCFYLMACPPFFLYLSRDRGYIQMTPLNKTKTA 60
Qy 77 LGFLMIVCMADLYFSWERSRGIFLAPVFLVSPTLIGITTLATFLIOLERRKGVOSG 136
Db 61 LGFLMIVCMADLYFSWERSRGIFLAPVFLVSPTLIGITTLATFLIOLERRKGVOSG 120
Qy 137 IMLFMWLVALCALAILRSKIMTALKEADAQVDFRDTFFVYFSLLIQVLSCFSDRSP 196
Db 121 IMLFMWLVALCALAILRSKIMTALKEADAQVDFRDTFFVYFSLLIQVLSCFSDRSP 180
Qy 197 LPSETHIDNPCEPSSASFLSRITFMWITGLIYNGIQOPLSGSLMSLNKEDTSEOVVPV 256
Db 181 LPSETHIDNPCEPSSASFLSRITFMWITGLIYNGIQOPLSGSLMSLNKEDTSEOVVPV 240
Qy 257 LVKMKKCEKATROPKPVYVSSKDPAPKSSKYDANEVEALIVKSPKEMNPSLPKV 316
Db 241 LVKMKKCEKATROPKPVYVSSKDPAPKSSKYDANEVEALIVKSPKEMNPSLPKV 300
Qy 317 LYKTFGPFLMSPEPKAIHDLMPGPOILKLIKFNVDYKAPDMOGYFTVLLFVTACL 376
Db 301 LYKTFGPFLMSPEPKAIHDLMPGPOILKLIKFNVDYKAPDMOGYFTVLLFVTACL 360
Qy 377 QTLVLAHQYFHCIVSGNRKIKTAVIGAVYRKALVITNSARKSSTVGEIVNLMSVDAOREMD 436
Db 361 QTLVLAHQYFHCIVSGNRKIKTAVIGAVYRKALVITNSARKSSTVGEIVNLMSVDAOREMD 420
Qy 437 LATYNNMTWSA.PIQVITIALYVLMNLGSPYAGAVWVWVAVNANAMKTKTQVAMK 496
Db 421 LATYNNMTWSA.PIQVITIALYVLMNLGSPYAGAVWVWVAVNANAMKTKTQVAMK 480
Qy 497 SKDNRKIKLMEINLINGIKVLKYLAMELAFKDKVLAIRBELVYLLKSAVLSAVGTFTWCT 556
Db 481 SKDNRKIKLMEINLINGIKVLKYLAMELAFKDKVLAIRBELVYLLKSAVLSAVGTFTWCT 540
Qy 557 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSALFNILRPPLNILPMVSISSIVQASVSKRL 616
Db 541 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSALFNILRPPLNILPMVSISSIVQASVSKRL 600
Qy 617 RIFLSHELEPDSIERRPVKDGGGNSITVNNATFTWARSPTLNGITFSIPGALVAV 676
Db 601 RIFLSHELEPDSIERRPVKDGGGNSITVNNATFTWARSPTLNGITFSIPGALVAV 660
Qy 677 VGVGCGKSSLSLALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGCQLEBPY 736
Db 661 VGVGCGKSSLSLALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGCQLEBPY 720
Qy 737 YRSYIQAACALLPDLIELIPSGDRTEIGEKNVLSGGQKORVSLARAAYSNADIYFDDPLS 796
Db 721 YRSYIQAACALLPDLIELIPSGDRTEIGEKNVLSGGQKORVSLARAAYSNADIYFDDPLS 780

Db 721 YRSVIAQACALLPDLLEILPSGDRTEIGEGVNLGGGOKORVSLADAVYSNADIVLFDDPLS 780
Qy 797 AVDAHVGHKIPENYIGPKMLKNTKRLIVTHSMGYLPQVDVITVMSGGKISEMGSYOELL 856
Db 781 AVDAHVGHKIPENYIGPKMLKNTKRLIVTHSMGYLPQVDVITVMSGGKISEMGSYOELL 840
Qy 857 ARDGAFAEFLRTVASTEOBDAENGVTGVSQPKGAKEKOMENGMVLVDSAGKOLQROUSS 916
Db 841 ARDGAFAEFLRTVASTEOBDAENGVTGVSQPKGAKEKOMENGMVLVDSAGKOLQROUSS 900
Qy 917 SSSYSIGISSRHNHSTAELOKAKEKETWKLMEADKAQTGOVKLSVTYWDYKAIQGLPISF 976
Db 901 SSSYSIGISSRHNHSTAELOKAKEKETWKLMEADKAQTGOVKLSVTYWDYKAIQGLPISF 960
Qy 977 LSTFLPMCHVSAASNYWLSLWTDDEPIVNGTOEHTVRLSVYALGISQIAIFGYSMA 1036
Db 961 LSTFLPMCHVSAASNYWLSLWTDDEPIVNGTOEHTVRLSVYALGISQIAIFGYSMA 1020
Qy 1037 VSIGGILASRCLHVDLHSLIRSPMSFEPTPSGNLVNRFSEKIDVTDSMTPEVTKMFMG 1096
Db 1021 VSIGGILASRCLHVDLHSLIRSPMSFEPTPSGNLVNRFSEKIDVTDSMTPEVTKMFMG 1080
Qy 1097 SLENVIGACIVILATPIAIIIPPLGLIYEFVQRFYVASSRQKRLSEVSRSFVSHFN 1156
Db 1081 SLENVIGACIVILATPIAIIIPPLGLIYEFVQRFYVASSRQKRLSEVSRSFVSHFN 1140
Qy 1157 ETLIGSVIRAFEBEERFIHQSDLKVDENOKAYPSIVANRWMLAVRLCEVCNCIVLPAAL 1216
Db 1141 ETLIGSVIRAFEBEERFIHQSDLKVDENOKAYPSIVANRWMLAVRLCEVCNCIVLPAAL 1200
Qy 1217 FAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVAVEBLKEXSETEKAPW 1276
Db 1201 FAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVAVEBLKEXSETEKAPW 1260
Qy 1277 QIOETAPSSWPQYGRVFPNNYCLRYREDLFPVLRHINVTINGEKVYIGVGTGAKSLS 1336
Db 1261 QIOETAPSSWPQYGRVFPNNYCLRYREDLFPVLRHINVTINGEKVYIGVGTGAKSLS 1320
Qy 1337 TIGLFRINESAEGIIIDGINIAKIGLHDLRFKTTIPQDPVLFSSGLRNNLDPFSQYSD 1396
Db 1321 TIGLFRINESAEGIIIDGINIAKIGLHDLRFKTTIPQDPVLFSSGLRNNLDPFSQYSD 1380
Qy 1397 BEWVTSLELHKLDPFVSALPDKLDHECAEGENLSVGOROLVCLARALLRKTKLIVDEA 1456
Db 1381 BEWVTSLELHKLDPFVSALPDKLDHECAEGENLSVGOROLVCLARALLRKTKLIVDEA 1440
Qy 1457 TAAVDETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1516
Db 1441 TAAVDETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1500
Qy 1517 QRGIFYSMADAGLV 1531
Db 1501 QRGIFYSMADAGLV 1515

RESULT 13

US-10-408-765A-1718
; Sequence 1718, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Martnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408.765A
; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1718
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1718
Query Match 71.1%; Score 7037; DB 4; Length 1388;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;
Qy 67 MTPLNKTKTALGFLMWIVCMADLFYSFWEBSRGIFLAPVFLVSEFTLIGITTLATPFIQL 126
Db 1 MTPLNKTKTALGFLMWIVCMADLFYSFWEBSRGIFLAPVFLVSEFTLIGITTLATPFIQL 60
Qy 127 ERKGVSSGIMLTFMLVALVCAIIRSKIMTALKEADOVLFRTITFYVYSSLLLIQL 186
Db 61 ERKGVSSGIMLTFMLVALVCAIIRSKIMTALKEADOVLFRTITFYVYSSLLLIQL 97
Qy 187 VLSGFSRSLPSETTHDPPCPRESSASFLSRTTFMWITGLIVRGYROPLEGSDMLSK 246
Db 98 -----NPCSSASFLSRTTFMWITGLIVRGYROPLEGSDMLSK 138
Qy 247 EDTSEQVVPVLVKNWKKCAKTRKQPVKYVSSKDPAPQKSSKVDANEVEALIVSPQ 306
Db 139 EDTSEQVVPVLVKNWKKCAKTRKQPVKYVSSKDPAPQKSSKVDANEVEALIVSPQ 198
Qy 307 KEWNPSEFKVLYTFCGYFLMSFPFKAIHDLMPSPQIIKLKIFVNDTKAPDMQGYFY 366
Db 199 KEWNPSEFKVLYTFCGYFLMSFPFKAIHDLMPSPQIIKLKIFVNDTKAPDMQGYFY 258
Qy 367 TULLFTYACIQTVLHOYFHICVSGMRITAVTGANRRALVITNSARSSVTEGIVNL 426
Db 259 TULLFTYACIQTVLHOYFHICVSGMRITAVTGANRRALVITNSARSSVTEGIVNL 318
Qy 427 MSYDAQRFMDLATYINMIWAPLOVILATYLMNLGSPSYLAGVAVWLVVPAVNAWAMK 486
Db 319 MSYDAQRFMDLATYINMIWAPLOVILATYLMNLGSPSYLAGVAVWLVVPAVNAWAMK 378
Qy 487 TKTYVAHMKSKONRIKLAMEIINGIKVLKYAMEIAFKDKVLAIRQEBLKVLEKSYLS 546
Db 379 TKTYVAHMKSKONRIKLAMEIINGIKVLKYAMEIAFKDKVLAIRQEBLKVLEKSYLS 438
Qy 547 AVGTFTWVCTPRIVALTFAVYVTTIDENNTLDQTAQFVSLATNIIARFPLNIPMTYSSI 606
Db 439 AVGTFTWVCTPRIVALTFAVYVTTIDENNTLDQTAQFVSLATNIIARFPLNIPMTYSSI 498
Qy 607 VQASVSLKRLRIFLSHELEPDSIERRPVQDGGTNSITVRNATFTWARSDDPTLNGITF 666
Db 499 VQASVSLKRLRIFLSHELEPDSIERRPVQDGGTNSITVRNATFTWARSDDPTLNGITF 558
Qy 667 SIEGALVAVVGVGCGKSSLSALLAEMDKVGHAIKGSVAVVPOQAMVIONDSIRENI 726
Db 559 SIEGALVAVVGVGCGKSSLSALLAEMDKVGHAIKGSVAVVPOQAMVIONDSIRENI 618
Qy 727 LFGCOLEBPYRNVIOACALLPDLLEILPSGDRTEIGEGVNLGGGOKORVSLARAYYSSNA 786
Db 619 LFGCOLEBPYRNVIOACALLPDLLEILPSGDRTEIGEGVNLGGGOKORVSLARAYYSSNA 678
Qy 787 DILFPDPLSAVDAHVGKHFENVIGPKMLKNTKRLIVTHSMGYLPQVDVITVMSGGKI 846
Db 679 DILFPDPLSAVDAHVGKHFENVIGPKMLKNTKRLIVTHSMGYLPQVDVITVMSGGKI 738
Qy 847 SEMGSYOELLARDGAFAEFLRTVASTEOBDAENGVTGVSQPKGAKEKOMENGMVLVDSA 906
Db 739 SEMGSYOELLARDGAFAEFLRTVASTEOBDAENGVTGVSQPKGAKEKOMENGMVLVDSA 798
Qy 907 GKLOLOLSSSSSYSDISHHNSTAELOKAKEKETWKLMEADKAQTGOVKLSVTYWDY 966
Db 799 GKLOLOLSSSSSYSDISHHNSTAELOKAKEKETWKLMEADKAQTGOVKLSVTYWDY 858
Qy 967 MKAIGHFISFLSIFLPMCHVSAASNYWLSLWTDDEPIVNGTOEHTVRLSVYALGISQ 1026
Db 967 MKAIGHFISFLSIFLPMCHVSAASNYWLSLWTDDEPIVNGTOEHTVRLSVYALGISQ 1026

Db 859 MKAIGLFIISLITLFCMCHVSALASNYMLSLMTDDPIVNGTOEHTKRLSVGALGISQ 918
Qy 1027 GIAVFGSMASVIGIGILASRCIAHYDLHSILRSPMSFFERTPSGNLVNRFSEKELDTVDSM 1086
Db 919 GIAVFGSMASVIGIGILASRCIAHYDLHSILRSPMSFFERTPSGNLVNRFSEKELDTVDSM 978
Qy 1087 IPEYIKNPMGSLFENVIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSRQLKRLSEV 1146
Db 979 IPEYIKNPMGSLFENVIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSRQLKRLSEV 1038
Qy 1147 SRSVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIYANMFLAVRLECV 1206
Db 1039 SRSVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIYANMFLAVRLECV 1098
Qy 1207 GNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIIVAVERLKE 1266
Db 1099 GNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIIVAVERLKE 1158
Qy 1267 YSETEKEAPMOIQETAPPSWPQVGRVFPNNYCLRYREDDLPVLRIHVVTTNGEKGIV 1326
Db 1159 YSETEKEAPMOIQETAPPSWPQVGRVFPNNYCLRYREDDLPVLRIHVVTTNGEKGIV 1218
Qy 1327 GRTAGKSSLTGLFRINESAGEIIDGINIAKIGLHDLRFKTIIPDPVLSGSLRM 1386
Db 1219 GRTAGKSSLTGLFRINESAGEIIDGINIAKIGLHDLRFKTIIPDPVLSGSLRM 1278
Qy 1387 NLDPFSQVSDDEWVTSLELAHLKDFVSALPKLDHECAEGENLSVQGRQLVCLARALLR 1446
Db 1279 NLDPFSQVSDDEWVTSLELAHLKDFVSALPKLDHECAEGENLSVQGRQLVCLARALLR 1338
Qy 1447 KTKILVLDEATAVADLETTDDLIQSTIRFOFEDCTVLTIAHRLNTIMDYTR 1496
Db 1339 KTKILVLDEATAVADLETTDDLIQSTIRFOFEDCTVLTIAHRLNTIMDYTR 1388

RESULT 14
US-09-939-853A-88
; Sequence 88, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 1528
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-88

Query Match 70.7%; Score 7002.5; DB 3; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

Qy 1 MALNGFCAGDSGDDPMDNNVNTWNTNSNPDFTKCFQNTVIVWPCFYLAACPPFYLYLSRH 60
Db 1 MALNSFCAGDSGDDPMDNNVNTWNTNSNPDFTKCFQNTVIVWPCFYLAACPPFYLYLSRH 60
Qy 61 DRGYQWMLNKTKATGALGFLMLIVCMADLFVSMFRSGIFLAPVFLVSPITLLGITTLA 120
Db 61 DRGYQWMLNKTKATGALGFLMLITICWADLFYSFMSRQGVLAAPVLLVSPITLLGITTLA 120

Qy 121 TFLIQLERRKGVSGSGIMLTFMWLVALYCALAALIRSKMTALKEDAOVDLFRDITFYVYS 180
Db 121 TFLIQLERRKGVSGSGIMLTFMWLVALYCALAALIRSKMTALKEDAOVDLFRDITFYVYS 180
Qy 181 LLLIQQLSCGSDSPLEFSETIHDNPNCPSSASFLSRITFMWITGLIVGYRQPLGSD 240
Db 181 LVLQVLASCFSDSPLEFSETIHDNPNCPSSASFLSRITFMWITGLIVGYRQPLGSD 240
Qy 241 LMSLNKEDTSEQVPLVVKMKCECAKTRKQPVKNVYS-SKDPAPQRESSKVDANEVEYA 299
Db 241 LMSLNKEDTSEQVPLVVKMKCECAKTRKQPVKNVYVAPRDPBPKPGSSQGLDVNEVEYA 300
Qy 300 LTVSPQKMNPSLEFKLYTTPGPFYLMSPFKAIHDLMPSCQIILKLIKPNDTKAP 359
Db 301 LTVSPQKMNPSLEFKLYTTPGPFYLMSPFKAIHDLMPSCQIILKLIKPNDTKAP 360
Qy 360 DMQGYFTYLLFVSACIQTLALHOYFHCIVSGKRITAVIGAYYRKAIVTNSARKST 419
Db 361 DMQGYFTYLLFVSACIQTLALHOYFHCIVSGKRITAVIGAYYRKAIVTNSARKST 420
Qy 420 VGEIYNLMSYDAQRFMDLATYINMWSAPLOVILATYLLMLNLGPSYLAGVAVVLMV 479
Db 421 VGEIYNLMSYDAQRFMDLATYINMWSAPLOVILATYLLMLNLGPSYLAGVAVVLMV 480
Qy 480 NAYMAKTKTYQVAHMSKONRIKLANEILNGIVYKLYAMELAFKDKVLAIRQELKVL 539
Db 481 NAYMAKTKTYQVAHMSKONRIKLANEILNGIVYKLYAMELAFKDKVLAIRQELKVL 540
Qy 540 KKSAYLSANGFTVWCPLVALCTPAVYVITDENNITLDQTAVALPILILFPNITL 599
Db 541 KKSAYLSANGFTVWCPLVALCTPAVYVITDENNITLDQTAVALPILILFPNITL 600
Qy 600 PMVTSIVQASVSLKRLRIFLSHELEPDSIERRPVQDGGTNSITVYNAFTWASDPP 659
Db 601 PMVTSIVQASVSLKRLRIFLSHELEPDSIERRPVQDGGTNSITVYNAFTWASDPP 659
Qy 660 TLNGITTSIPGALVAVVGQVCGSKSLSLALAEMLRVSGHVAIKSVAVVPOQAMION 719
Db 660 TLNGITTSIPGALVAVVGQVCGSKSLSLALAEMLRVSGHVAIKSVAVVPOQAMION 719
Qy 720 DSIARENILFGQCEPPTYSYVIOACALLPDEILIPSDRETEIGKGNLSGCGORVSLA 779
Db 720 DSIARENILFGQCEPPTYSYVIOACALLPDEILIPSDRETEIGKGNLSGCGORVSLA 779
Qy 780 RAYVSNADYILFDDPLSAVDAAHVGKHIIFENVYIGPKGMLKNKTRILVTHMSYLPQVDYI 839
Db 780 RAYVSNADYILFDDPLSAVDAAHVGKHIIFENVYIGPKGMLKNKTRILVTHMSYLPQVDYI 839
Qy 840 VMSGKISSEMGSYQELLARDGAFAPLRTYASTEQEODAEENGVTGVSGPKAKOMENG 899
Db 840 VMSGKISSEMGSYQELLARDGAFAPLRTYASTEQEODAEENGVTGVSGPKAKOMENG 896
Qy 900 MLVYDSAGKQLOROLSSSSYSGDISRHNSTALQKAEKERTYKLMEDKAOQGVK 959
Db 900 MLVYDSAGKQLOROLSSSSYSGDISRHNSTALQKAEKERTYKLMEDKAOQGVK 955
Qy 960 LSVYMDVYKATGLFISFISFLFCMCHVSALASNYMLSLMTDD-PIYNGTOEHTKRLSV 1018
Db 960 LSVYMDVYKATGLFISFISFLFCMCHVSALASNYMLSLMTDD-PIYNGTOEHTKRLSV 1015
Qy 1016 YGALGILQGAALIFGYSMASVIGIGIFASRRRLHDLVYVLRSPMSFFERTPSGNLVNRFSSK 1075
Db 1016 YGALGILQGAALIFGYSMASVIGIGIFASRRRLHDLVYVLRSPMSFFERTPSGNLVNRFSSK 1075
Qy 1079 ELDTVDNMTPEYIKPMGSLFENVIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSR 1138
Db 1079 ELDTVDNMTPEYIKPMGSLFENVIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSR 1135
Qy 1139 QKRLSESVSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIYANRW 1198
Db 1136 QKRLSESVSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIYANRW 1195
Qy 1199 LAVRLCEVGNCIYVFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNI 1258

Db 1196 LAAVLECGNCIVLFAALFAVISRHSLSAGLVSVSLQITAVIMLVMSSEMETNI 1255
Qy 1259 VAVERLKEYSETEKEAPWQIOETAPSSWPPOVGRVEFPNNCLARRREDLDFVLRHINTIN 1318
Db 1256 VAVERLKEYSETEKEAPWQIOETAPSPWPSHGRVEFPDYCARREDLDLVLRKHINTIE 1315
Qy 1219 GGEVGVIGRTGAGKSSLTLLGFRNESAGEIITIDGINIAKIGHDLRPFKITTIIPODPV 1378
Db 1216 GGEVGVIGRTGAGKSSLTLLGFRNESAGEIITIDGINIAKIGHDLRPFKITTIIPODPV 1375
Qy 1379 LFGSSLRNNLDPFQOYDEEYVWLSLELAHLKDQVSALEPDKLDHECAGGENLSVGOQOLV 1438
Db 1376 LFGSSLRNNLDPFQOYDEEYVWMALELAHLKGFSALPDKANHECAGGENLSVGOQOLV 1435
Qy 1439 CLARALLRKIKYILVDEXTAAVLETDLLIOSTIRTOGEDCTVTIAHRLNTINDYTRVI 1498
Db 1436 CLARALLRKIKYILVDEXTAAVLETDLLIOSTIRTOGEDCTVTIAHRLNTINDYTRVI 1495
Qy 1499 VLDKGETOEVGAPSDLLQORGLFYSMADAGLV 1531
Db 1496 VLDKGEVRECGAPSELLQORGIIFYSMADAGLV 1528

RESULT 15
US-10-618-281-63
; Sequence 63, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618, 281
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIORITY FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 1303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-281-63

Query Match 61.3%; Score 6068.5; DB 5; Length 1303;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1229; Conservative 17; Mismatches 57; Indels 207; Gaps 10;

Qy 22 WNTSNPFTKCFQNTVWVWPCFYLMACFPFYFLYLSRHDGYIOMTPINKTKTALGFL 81
Db 1 WNTSNPFTKCFQNTVWVWPCFYLMACFPFYFLYLSRHDGYIOMTPINKTKTALGFL 60
Qy 82 WIYCMADLFYSFEMRSRGIFLAPVFLVSPILLGITLLATFLQLERRKGVQSSGIMLTF 141
Db 61 WIYCMADLFYSFEMRSRGIFLAPVFLVSPILLGITLLATFLQLERRKGVQSSGIMLTF 120
Qy 142 WLVVLVCAALILRKIKITALKEDQVNDLFRDITFYVYFSLLLIQVLSCSFDRSPLFSET 201
Db 121 WLVVLVCAALILRKIKITALKEDQVNDLFRDITFYVYFSLLLIQVLSCSFDRSPLFSET 177
Qy 202 IHDPNPCESSASFLSRTITFMWITGLIVRGYQPLEGSDLMSLNKEDTSEQVVPVLVKNW 261
Db 178 IHDPNPCESSASFLSRTITFMWITGLIVRGYQPLEGSDLMSLNKEDTSEQVVPVLVKNW 237
Qy 262 KKECAKTRKQPKVYVYSKDPQPKESKVDANEVEALLVYSQKEKNSPLFVLYKTF 321
Db 238 KKECAKTRN-----SSGSGESCSANTEALF-----PA--PTCHKSF 271
Qy 322 GPVFLMSFPFKAIHDLMMFSGPQILKLLIKFVNDTKAPDWQGYFYTVLLPFTACLQTLVL 381

Db 272 QALSLL-----LCRLIKFVNDTKAPDWQGYFYTVLLPFTACLQTLVL 314
Qy 382 HQYHIFGVSGMRIKTAVIGAVYRKALVTINSARKSITGEIVNLMSVDAQRPMDLATYI 441
Db 315 HQYHIFGVSGMRIKTAVIGAVYRKALVTINSARKSITGEIVNLMSVDAQRPMDLATYI 374
Qy 442 NMITSAPLOVTLALYLLMLNLGSPVLAGVAVMLVMPVNAVMARKTQYVAHKKSKNR 501
Db 375 NMITSAPLOVTLALYLLMLVAPDLTAV-----SSKAHKKSKNR 416
Qy 502 IKLNLNELLNGIKVLAELAFQOKYLAIRQBELKYLKKSAYLSAVGTFTWVCCPLVA 561
Db 417 IKLNLNELLNGIKVLAELAFQOKYLAIRQBELKYLKKSAYLSAVGTFTWVCCPLVA 476
Qy 562 LCTFAVYVITDENNILDQAFAVSLALENLIRPELNLPMVSISSIVQASVLSKRLRIFLS 621
Db 477 LCTFAVYVITDENNILDQAFAVSLALENLIRPELNLPMVSISSIVQASVLSKRLRIFLS 528
Qy 622 HEELEPDSIERRPYKDGCGTNSITVRNATFTWASDDEPTLNGITFSPICEALVAVGQVG 681
Db 529 GATSERGPWGSFRKQG-----TROASPSVABGYLCRESITFSPICEALVAVGQVG 581
Qy 682 CGKSLSLAALLAEMDKVEGHVAIKGSVAVYVPOQAWIONDSLRENILRGCOLLEBPYRSVI 741
Db 582 CGKSLSLAALLAEMDKVEGHVAIKGSVAVYVPOQAWIONDSLRENILRGCOLLEBPYRSVI 641
Qy 742 QACALLPDLLELPSGDRTEIGEKVNLSGQKQKQVSLARAVSNADYLPDPLSAVDAN 801
Db 642 QACALLPDLLELPSGDRTEIGEKVNLSGQKQKQVSLARAVSNADYLPDPLSAVDAN 701
Qy 802 VGKHPENYVIGPKMLKXKTRILVTHSMSTLPQVDYIIVMSGKISENGSYQBLIANDGA 861
Db 702 VGKHPENYVIGPKMLKXKCL----- 724
Qy 862 FAELFRYASTBEOQDAEBNGVGVSGPGEAKOMENGMVTDASGQOLQRLSSSSSV 921
Db 725 -----S 725
Qy 922 GDLSRHNSYAELOKAEKKEETWKLMEADKAQOGYKLSVYMDYMAIGLFIISLIFL 981
Db 726 CDL-----QYKLSYMDYMAIGLFIISLIFL 753
Qy 982 FMCNHSALASNYWLSMTDDPIVNGTOEHTKQVLSYVAGLIGISOGIAVFGYSAVSIIG 1041
Db 754 FMCNHSALASNYWLSMTDDPIVNGTOEHTKQVLSYVAGLIGISOGIAVFGYSAVSIIG 813
Qy 1042 ILSRCLHVDLHSLISPSFPERTPSGVNLNRFSEKELDTVSMITEVFKMFGSLFNV 1101
Db 814 ILSRCLHVDLHSLISPSFPERTPSGVNLNRFSEKELDTVSMITEVFKMFGSLFNV 873
Qy 1102 IGACIVILLATPLAIIIPPLGILYFPVQFVYVASSQLRLLESRSRPVSHNETLIG 1161
Db 874 IGACIVILLATPLAIIIPPLGILYFPVQFVYVASSQLRLLESRSRPVSHNETLIG 933
Qy 1162 VSVIRAFEBQERFHQSDLVYDENOKAYYSIVANRWLAVLECGVNCIVLPAALPAVIS 1221
Db 934 VSVIRAFEBQERFHQSDLVYDENOKAYYSIVANRWLAVLECGVNCIVLPAALPAVIS 993
Qy 1222 RHSLISAGLVGLSVSYLQVTTYLNMLVRMSSEMETNIIVAVBRLEYSETEKEAPWQIOET 1281
Db 994 RHSLISAGLVGLSVSYLQVTTYLNMLVRMSSEMETNIIVAVBRLEYSETEKEAPWQIOET 1053
Qy 1282 APPSPWQVGRVEFRNCLARRREDLDFVLRHINTINGEKGVGIVGTGAGKSSLTGLF 1341
Db 1054 APPSPWQVGRVEFRNCLARRREDLDFVLRHINTINGEKGVGIVGTGAGKSSLTGLF 1113
Qy 1342 RINESAGEIITIDGINIAKIGHDLRPFKITTIIPODPVLFGSGLRMMNDLPPSQYSDSEVMT 1401
Db 1114 RINESAGEIITIDGINIAKIGHDLRPFKITTIIPODPVLFGSGLRMMNDLPPSQYSDSEVMT 1173
Qy 1402 SLELAHLKDFVSLPDLDEHCAEGGENLSVGOQOLVCLARALLRKIKYILVDEXTAAV 1461

Db	1174	SLELAHLKDFVSALPDKLDHECAREGENLSVGOROLVCLARALRKTKIIVLDEATAVD	1233
QY	1462	LETDDLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQORGLF	1521
Db	1234	LETDDLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQORGLF	1293
QY	1522	YSMAKDAGLV	1531
Db	1294	YSMAKDAGLV	1303

Search completed: December 15, 2005, 15:46:47
 Job time : 169.832 secs

Qy	855	LLAADGAFAEFLRPTASTAEQODAEENVTGVSGGKXAKOMENMLVTBDAK--QUR	912
Db	360	LLQRKGAIVCL-----DQARQPDGR-EGTEFGTSTKPDG-----TSAGRRPELIR	407
Qy	913	QLSSSSVSGGISRHHNSTAEIQRKAKEET----MKLEADKAQTQGVKLSTVYWDYMK	968
Db	408	ERSIKS-----VPEKDRITSEQ-TEVYLLDDDRDRAWGAK-DSIQYGRVATYHIAVLR	460
Qy	969	ALGFISFLSIFLFWCNVNSALASNYMLSLWTDDEIVNGTOEHTKRLSYVGALGISGCI	1028
Db	461	AVGPTPLCYALFLFLCQOVASFCRGWYMLSMADPADVAGGQCTQALRFGIGIFGLGCLQAI	520
Qy	1029	AVFGSMAVSYGGILASRCLAVDLHSILRSMSEFEETPGSNLYNRSEKELDTYDSNIP	1088
Db	521	GLFASMAVLLGGARASRLFORLLMDVARSPISEFEETPIGHLNRRSEKELDTVDVDIP	580
Qy	1149	SPVYSHFNETLLGVSVIDAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVN	1208
Db	641	SSVCSHMAETFGQSTVVRARFTQAPDFVAQNNAVRDESQRISPPRLVADRMLAANVELLGN	700
Qy	1209	CIVLFAALFAYISRSLSAGLVGLSVSYSLQTTYTLNMLYNRSMENINIVAVELKXYS	1266
Db	701	GLVFALAAACAVLSKXKHSAGLVGFSVSALQYTQTLQWVRNWTLENSIVSVERMODYA	760
Qy	1269	ETEKEAPNOIGETAPSSWPQGVAREFNKYCLRYREDDLVFVRHINVTINGEKGIYGR	1328
Db	761	WTPKEAPNRLPFCNAQPPMFGQGI-EFRDFGIRYRPELPLAVQVYSFKIHAGEKGIYGR	820
Qy	1329	TGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPODVLFSSGLRML	1388
Db	821	TGAGKSSLASGLRLRQEAABEGTIWDGVPYIAVGHTRLSRSITIIPODPIFPGLRML	880
Qy	1389	DPFSQYSDEEVTSLAHLKDFVSLPDKLHCEKAGEGNSVGRQLVCLARLLKRT	1444
Db	881	DLOEHSDBEAIYAALETETVLKALVSLPQOLQYCKADREEDLSVOKOQLCLARLLKRT	940
Qy	1449	KIVLVDEAPTAAVDETDDLIQSTIRPOFEDCTVLTAHRLNIMDYTRIVYLDKEIOEY	1508
Db	941	QILIIDERTAAVDPGTEILOQMAMLSWFAQCTVLLIAHRLRSVMDCARVAVMDKQVABS	1000
Qy	1509	GAPSDLLQORGLFYSMAKDAGLV	1531
Db	1001	GSPAQLLAQKGLFYRLAQESGLV	1023
RESULT 2			
US-11-090-439-24			
Sequence 24, Application US/11090439			
Publication No. US20050266442A1			
GENERAL INFORMATION:			
APPLICANT: Squillace, Rachel			
APPLICANT: Weiner, Michael P.			
TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null			
FILE REFERENCE: 24318-502			
CURRENT APPLICATION NUMBER: US/11/090,439			
CURRENT FILING DATE: 2005-03-25			
PRIOR APPLICATION NUMBER: 60/556,344			
PRIOR FILING DATE: 2004-03-25			
NUMBER OF SEQ ID NOS: 62			
SOFTWARE: PatentIn Ver. 2.1.1			
SEQ ID NO 24			
LENGTH: 1581			
TYPE: PRT			
ORGANISM: Homo sapiens			
US-11-090-439-24			
Query Match	21.4%	Score 2115;	DB 7; Length 1581;
Best Local Similarity	31.9%;	Pred No. 3e-147;	

Matches	527;	Conservative	305;	Mismatches	564;	Indels	256;	Gaps	37;
QY	32	CFONTVLVWPCFLYMACFPPEFL	-YLSRHDGYIQMTPLANKTALGF---	LLMTVCM	86				
Db	26	CFVDALNVVPHVFLFTFTFFLFI	FGMSOSKAIHHS-----	TWLFPHGNLRMTLTF	79				
QY	87	ADLVYSFWEBSRGFLFAPV----	FLVSPF----	LLGITTLATFLIQERRGVSSG	136				
Db	80	MLLFVLVCEIAEGILSDGVESH	HLHLYMPGMAFMAAVTSV--	YVHNLETSNFPKLLI	137				
QY	137	IMTFMWALVACALAIRSKIMTAL	KEDADVDFRDTTF-----	-VYESLILQVLVS	189				
Db	138	ALIVYMLTALF-----	TKTKIKVKFLDHAIGSQARF	CLTGVLIVLIGWLLIVE--	188				
QY	190	CFSDRSPLESETIHDNCPSES-----	SASFLSRITFWITGLIVRGYRPLEG		238				
Db	189	VIRVRRYFFKTPREVKP-PED	LODLGRLFPQPVNLISKGT	YMMNNAFITAHKAKYI--	245				
QY	239	SDMSLKNEDISEQVNVPLVK----	NMKKECAKTRKQVKKVYSSK	DAQPRRESSKYDANE	295				
Db	246	-DLRAIG-----	LPLAMALTNVORLC-----	-E	268				
QY	296	EVEALIVKSPQ-KEMNSSLFV	LKXTGPEYFLMSEFFKAI	HDLMMSGPQIILKLKIV-	353				
Db	269	AFDAQVDIOGTGGAIRALIM	ALSHAGRLVLSSTRILLAD	LIGFAPGLCITFGIVHLG	328				
QY	354	--NDTKAPDWQ-----	-GYFTVLLFVTAQIQTIVL	HOYFHIQVSGMRK	396				
Db	329	KENDVFOGKTOFLGVYFVS	SOEPLANNVVLALFLFL	ALLQRTFLQASYYAILEGTINLR	388				
QY	397	TAVIGAVYRKA--LVITNSARKS	TGEIVLMSVDQRFNDLAT	YINMLISAPLOVYLA	454				
Db	369	GAIQTKIYNKIMHLS	TNSLMSGEHTAQICL	VALIDTNQLMWPFPLCPNL	MAMVQIIVG	448			
QY	455	LYILMLNIGPSVLGVA	VMVLMVYNAMAKTKTYO	VAHMSKDNRIKILNEILNGIKV	514				
Db	449	VILKYILIGVSA	LGAIVYIILAAVQYFVAL	KLSQAQRSTLEYSNEHRLKQTNEMLRGKL	508				
QY	515	LKUYAMELAFKDYALIR	OEELKYLKKSAYLSAVGT	FTWCTPFLVALCTPAVVYTTIDEN	574				
Db	509	LKUYAMENI	PFTRVETTRRKEMTSLRA	PAIVTYSISIFENYTAI	PLAAVLITFVGHVSPFKE	568			
QY	575	NILDAQTFVSLAFNLIR	FLPNILPMTYSSIVQASVSL	KRLRILFLSHELEPSTI----	630				
Db	569	ADFSVSAFASISL	PHILVLPFLFLSSVRS	TVALVSVOXLSEFLS	AAELREQCAPH	628			
QY	631	-----ERRPVK-----	-----	-DGGGTV-SITVNA	AFPTM	653			
Db	629	PTPGSPASKYQAVPLRV	VNNRKRPRABEDCRGLT	PGLOSLVPSADGADADNCC	YQINGCYFTW	688			
QY	654	ARSDPPTINGITFSI	PEGALVAVVGVQVCGKSS	LLSALLAEMDKVEG-----		700			
Db	669	TPDGIFTLSTNITIR	IPRGQIMYGVQVCGKSS	ILLALGEMQKVASGVFMS	LPDSEIG	748			
QY	701	-----HVAIKGSVA	YVYPQQAIONDSLRENIL	IFGCQLEEPYRSYIQACLL	747				
Db	749	EDPSERETATDLDIR	KRGPAVAYASQKFWL	NATVEENIIIESEFPNK	RYKVIIEACSLQ	808			
QY	748	PDLFLPESGDTEI	GEKVNLSGGOKQVSLAR	AVSADYVLPDDPL	ISAYDAVKGRI	807			
Db	809	PDIDLIPGDDTO	IGERGINISGGORQRI	SVARALYQHANVAF	LPDDPSALDIHLSHLM	868			
QY	808	ENVISPKMUK-----	NKTRILVTHSMYSYL	PQOVVYIYMSGAKIS	SEMSSYOELARDAF	862			
Db	869	Q-----AILELRL	DDKRTVVLVTHKQLQ	PHADMIIIMKQGT	IQREBTLADPQRS	923			
QY	863	AELFTVYASTBO	DAENGVTVGSGPGKEAK	KOMENGMVTD	SAGKOLQROLSSSSSYSG	922			
Db	924	FEHMKTIMN-RQDQ	LEKEIVT-----	ERKATEPQG-----	LSRAMSSPDGL--	965			
QY	923	DISRHHNSTALQ	RAKKEETWKLMEADKA-----	QYGVYKLSVYNYDMY	KALGLFISFL	977			
Db	966	-----LODEEBEE	BAEEDBNISMLHQAEL	IPWRA	CATYLSAGIILLSTL	1014			

Qy	978	SIPELMCHVGNALSNMYLSMTDDPIV-----NTOGHT-----KRLSYGALG	1023
Db	1015	LVFSQLEKHMVLVADIDYWLAKWTDSALTLTPAARNCSLSOECTDQTVYAVAVFYVLCSLG	1074
Qy	1024	ISOGIAVFGVMAVSIIGGIIASRCILVHDLHSILRSPMSFEPTPSGNLVNRFSELDTV	1083
Db	1075	I---VLCIVTSYVWMTGLKXAKRLHRLBLKRIILAPKRFETTPGLSILNRFSSDCNTI	1131
Qy	1084	DSMIEVIMKFMGSLFNIVIGACTIVLLATPIAIIIPPLGLIYFPVQRFYVASSRQLRL	1143
Db	1132	DQHPSTLECSRSRTGLCVSALAVISYTPVFLVALTELAIVCFYIOKXFPVASHDLOQL	1191
Qy	1144	ESVRSRPYYSHENETLLGVSIVIRAEFEERFIHOSGLKVDENOKAYYPSIYANRYLRL	1203
Db	1192	DDTYQLPULSHPAETVBGLTTTIRAFRYEARFOQKILEYTDSSNNIASLFLTANRWLEVRM	1251
Qy	1204	ECVGNCIYLFALPAFVVIS--RHSSAGLVGSVSYSLQVTTYLMLVLVMSSEMETNIYAV	1261
Db	1252	EYIGACVLLIAVVISISMSLHRELSAGLVGLTYALNAVSYLMMVGNLADMELOGLAV	1311
Qy	1262	ERLKEVSETEK-----APWQIOETAPSSMPQVGRVEFNNYCLRYREDIDPVLRLINV	1315
Db	1312	KRIHGLKTEAESYEGGLAPSLI-----PKWMPDQCKIQIQLSVRYDSSILKPYLKHVNA	1366
Qy	1316	TINGEKXGIVGRGAGKSSLTJLGRINESAEGIIIDGINIKIGIHDIRFKITTIPO	1375
Db	1367	LISPOQKIGICGRSGKSSPSLAFPRVNDTFEGHIIIDGIDIALPLHTIRSRSLILO	1426
Qy	1376	DPVLFSGSLRNMLDPFSGSDSEVWTSLELAHLKDIFYALPDKLDHCEAGEGENLSVGOR	1433
Db	1427	DPVLFSGTIRNLDPERKCSDSLTEWALEINQKLKVAGALLGGDAIITBEGGENFSGQOR	1466
Qy	1436	QLVCLARALLKTKILVIDEAPVAVDLETDLIOSTIRFOEDCTVLTIAHRLNTINDYT	1495
Db	1487	QLFCIARAFVAKTSLIFIMDEATASIDAMATENILQVWMTAARDRTVVVTIARVHTIISAD	1546
Qy	1496	RVLVLDKEGIOEYCAPSDILOOR-GLFPMXAK	1526
Db	1547	LVIYVKRGAILIEFDKPEKLLSRKDSVFASFVR	1578
RESULT 3			
US-10-995-561-527			
: Sequence 527, Application US/10995561			
: Publication No. US20050272054A1			
: GENERAL INFORMATION:			
: APPLICANT: CARGILL, Michele al.			
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH			
: TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF			
: TITLE OF INVENTION: DETECTION AND USES THEREOF			
: FILE REFERENCE: C0001559			
: CURRENT APPLICATION NUMBER: US/10/995,561			
: CURRENT FILING DATE: 2004-11-24			
: NUMBER OF SEQ ID NOS: 85702			
: SOFTWARE: FastSeq for Windows Version 4.0			
: SEQ ID NO 527			
: LENGTH: 1568			
: TYPE: FRT			
: ORGANISM: Homo sapiens			
: US-10-995-561-527			

	Query Match	21.2%;	Score 2101.5;	DB 6;	Length 1588;
	Best Local Similarity	31.8%;	Pred. No. 2.9e-146;		
	Matches 527;	Conservative 305;	Mismatches 564;	Indels 2c3;	Gaps 38;
QY	32 CFONTLVWVCEFLMACFPFYFL-YISRHDRTGIOMTPLNKTTALGF---- <td>:</td> <td>:</td> <td>:</td> <td>:</td>	:	:	:	:
	26 CFWALNNVPHAVFLFTFFLTFPIFLIGWGSSKVIHHNS-----TWLFHFGHNRLMILTF 79	:	:	:	:
Dh		:	:	:	:
OY	87 ADLYSFRERKRGLFVLAVY-----FLVSPT----LLGITTLATLPQLERRRGVSSG 136	:	:	:	:
Dh	80 MLPLVLVACEIAEGISDGVESHHLALPMPGMAFMVAIVSV--YYNIITSNPFKLII 137	:	:	:	:

QY	137	IMLTFMLVALVCALAILRSKIMTALKEDAOVDFRDTTF-----YVFSLLILQVLVS	169
Db	138	ALIVYMTLAI-----TTIKRFXFLDHAIGFSQIRFLTGLVILYGMELIVE--VN	188
QY	190	CFSDRSLPSETIHDHPNCSES-----SASFISITFPMWITGLVIRGROPLBG	238
Db	189	VIRVRKIFETKPREVKP-PEDJLQDGVRFLOPFPNLDSKGYTMMNAFIKTHAKFI--	245
QY	239	SDLMSLNKEPTEBOVAVLVK--NMKKECAKTRKOPVKVAVVSSKDPACPKSSKYDANE	295
Db	246	-DLRAIK-----LPIARBLATNVQRLC-----E	268
QY	296	EVEALIVKSPQ-KEMNSLFKVLKYTFGPFYFMSFPFKAIHIDLMFSGPOLIKLKEV-	353
Db	269	AFDAQVKDQIGTQGARAIWQALSHAFGRRLVSSFRILADLIGFAGPLCTIGVIDHIG	328
QY	354	--NDTKAPDMQ-----GYFTVLLFPTACQTLVLAQYHICVSGSRKIK	336
Db	329	KENDVFQPKTQFLGVYFVSSQSEFLANAYVLAVLLFLALLQRTFLQASVVALETGINLR	368
QY	397	TAVIGAYYRKA--LVITNSARKSSITGCEIVNLSMAQORFMDATATYIMMWSAPLOYILA	454
Db	389	GAIOTKTYKIMLSTNSLSMGEMTQOICNLVAIDTNQMLMFFFLCPNLMAHPVQIIVG	448
QY	445	LYLLMLNIGBSVLGAVAVMLAVPVNAVMAKTKTYQVAMHMSKDNRIKLMNEILNGIKV	514
Db	449	VILLYITLIGVSALIGAIVIIILAPVQYFVATKLSQARSTLEVSNERLKQTNMELNGIKV	508
QY	515	LKYAMLEAFKDKYLAIROELKVLKKSAYLSAVGFPTWCTPFLVALCTFAVYVITIDEN	574
Db	509	LKYAMENIRFRTVETTRRKEMTSLBAFALYTSISIFPMVAIPIAAVLITFQHVSEFFKX	568
QY	575	NILDAQFAFSLFENIRFENILRPMVLSITQVASYSLKRLAIFLSHEELBEDSI----	630
Db	569	ADPSPSVAFASLSLFHILTPFLSSVRSVTKALVSQQKSEPLSSAIREEQCAPIIE	628
QY	631	-----ERRPVK-----DGCGTN-SITYRNAFTFW	653
Db	629	PTQGRASKXQAVPLRVNRKRPARBEDCRGLTGPLOSIVPSADGDANDCVOQMGVFTW	668
QY	654	ARSDPPLNGITPSIPGALVAVVGVGCGKSSLSALLAEMDKVEG-----	700
Db	689	TPDGIFPLSNITIRIPGQLTMIIVGVGCGKSSLLAALGEMQKVGAFAVWSSLPDSIEG	748
QY	701	-----HVAIKGSVAAPVPOQAWIQNDSIAENILFGQULEEPPYRSYIQCALL	747
Db	749	EDSPSERETATDDIRKRGFVAASQKPMILNLTVEENIIFESFNKORIKWATECSAQ	808
QY	748	PDLIELPSGDRTEIGEGKVNLSGGQKORVSLAAVSNADIYLFDDPLSLAVDAHYKHIF	807
Db	809	PDIDILPHGQTOIGERGIVLSGGQKORISVABALYQHANVYVLDLDPFSALDHLSDHLM	868
QY	808	ENVIYGPCKMLK-----NKTIRILVTHSMYSYLPQVDVLIIVNSGKISIMSGSYOELIARDGAF	862
Db	869	Q-----AGILELRDOKRIVLVLTHTLKQYLPHADWIIIAKKDGTIQREBGLTKQORSECDL	923
QY	863	AELFRTYASTROGDAENGCVTSVSGPGEAKKMEMGMLVTBSAGQOLROULSSSSSYSG	922
Db	924	FEHMKTLMN-RQDOELEKETVLT-----ERKATEPPOG-----LSRAMSSRDL-	965
QY	923	DIRHNNSTAELOKAEKKKETKMLBADA-----QTGOVUKLSYVWDYMKAIIGLISFL	977
Db	966	-----LQDEEBEEBAASEBDDULSMHLQRAEITPWRACAKYLSGILLLSL	1014
QY	978	SIFLPMCNHVSALASVWYLSLMTDDIV-----NGTOEHT-----KVRLSVYVGLG	1023
Db	1015	LVFSQILKMHVLAIVADYMLAKMTDSALTITLPARANCSLSQECTLQDTQVYAMVFTVLSIG	1074
QY	1024	ISGCIAYFGYMAVVSIGGILASFCLHVDLHSLTSLRPMSPFEFTSPGCVLNRFSKELDTV	1083
Db	1075	I---VLCLVTSVYVENVETGLKVARRHRSILNRIILAIPMFEEFTTPIGSSILNRSSCNTI	1131

QY	923	DISHNHNSIABEVLGAELNANDBEINMCHALDA	-----	QIGGVKSNVSVNMLIRKALSGU	1131
Db	966	-----	-----	-----	-----
QY	978	SIFELMCNHNVSALASNNVLSIMTDDPIV	-----	NGAQEHT	1023
Db	1015	LVSFQLKHNVLVAIDYWLAKMTDSALTILTPARNCSLSQECTLDQTYVAMVFTVCSIG	1074		
QY	1024	ISOGIAVFGYSMAVSIGGILASFCHVDLHSLTILRSPMSGPFERTSPGSLVNRPSKELDV	1083		
Db	1075	I---VLCVLVSIVVENVTGLKVARLRHRSILNRIILA PMEFETPTPGSLINRSPSCNNT	1131		

QY 1084 DSMIPEVTKMFMGSLFNIVIGACIVILLATPIAIIIPPLGIYFVQRFYVASSRQKRL 1143
DQ 1132 DQHISTTECLSRSTLLCVSALAVISYVTPVFLVALPLAIVCYFIQCFRVASRDLOQL 1191
QY 1144 ESVERSPVYSHENFTLGVSVIRAFEBQERFIHQSDLKVDENQAKAYYSIVANRWLAVRL 1203
DQ 1192 DDTQPLSLSHFAETVEGLTTIRAFRYEARFOQKLEVTDSNNIASLFLTANRWLEVRM 1251
QY 1204 -----ECVNGCIVLPALPAFVVIS--RHSLAGIVGSVSYSLQVTTYLNLVMSSES 1254
DQ 1252 ATPPEQYIGACVYLAAVVISNSLHRELSAGIVGLTYALMVSITLMMWRNLADM 1311
QY 1255 ETNIVAEVRLKEYSETEKE-----APWQIETAPSSWPQVGRVEFRNYCLARYEDLD 1308
DQ 1312 ELQGVAVRIRHGLKTEAESYEGLLAPSLI-----PKWMPQKQIQIONLEVRSDSLKP 1366
QY 1309 VLKRIINTINGEKVIGVIRGTAGKSSLTGLFRINSEABEGEIIIDGINIKIGLHDIRF 1368
DQ 1367 VLKVNMLISPGQKIGICGRGSGKSPSLAFFRWVDFEGBHIIIDGIDAKLPLHTLSRS 1426
QY 1369 KITIIPODPVPLSGSLRMLNDPFSQYSDBEVWTSJELANLKDQFVSLPDKLDHECAEGE 1428
DQ 1427 RLSITILQDPVPLSGTIRNDPBERKCSSTIMEALEIQKLQVYKALPGGLDAITTEGGE 1466
QY 1429 NLVSGORQVCLARALLRKTKILVLDERTAAVDLETDDLQSTIRTOFEDCTVLTIAHRL 1468
DQ 1487 NFSOGORQPLCLARAFVAKTSIFIMDEATASIDMATENILQKVMTAFADRVTVTIAHRV 1546
QY 1489 NTIMDYTRVYVLDGGEIOEGYCAPSPDLOOR--GLFYSMAK 1526
DQ 1547 HTLSADLVIVLKGGAILEPDKPEKLSRKDSVFASFVR 1585

RESULT 4

US-11-090-439-26
; Sequence 26, Application US/11090439
; Publication No. US2005026642A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberosus Sclerosis Null
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 26
; LENGTH: 1581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-26

Query Match 21.2%; Score 2100; DB 7; Length 1581;
Best Local Similarity 31.7%; Pred. No. 3.7e-146;
Matches 524; Conservative 306; Mismatches 567; Indels 254; Gaps 37;

QY 32 CFQNTVAVWPCFYIMACFPFYFL-YLSRHDRGYIQMTPLNKTALGF---LIMIVCW 86
DQ 26 CFVALANVPHVFLIFITFPIIFIGMGSSQSKVHIHS-----TWLHPGHNLMWILTF 79
QY 87 ADLFYSFMRSGRTFLAV-----FLVSP-----LIGITTLAFLQLQERKQVQSSG 136
DQ 80 MLFLVLCVIAEGHISDGVTSNHLALTWPAGAMATVAVTVV--YHNHITSNFPKLLI 137
QY 137 IMLTFLVALVCALALRSKIMTALKEQVDFRDTF-----YVFSLLIQVLVS 189
DQ 138 ALIVVMTLAFI-----TKTIKFKVCLDHAIGFQSLRCLGLVILYGMILLVE--VN 188
QY 190 CFSDBSPFLFSET---IHDPNCPRESSASFL-----SRITFWMITGLIVRGYRPLBGS 239

DQ 189 VIRVRRYIFFKTPEVKKPREDLOLQVRFLOPFVNLPSKGTWMMNAPITKAHKPI--- 245
QY 240 DIMELNKEDTSEGVVPLVVK---NMKKECAKTRKQPVKVYVSKDPAQPRESSKYDANEE 296
DQ 246 DLRAIGK-----LPIYMAALTNVRLC----- 269
QY 297 VEALIVKSPQ--KEWNPSELKVLKYKTGPYFLMSEFFPAIHDLMFSGPQILKLIV-- 353
DQ 270 FDAQVRKDIOGTQCARAIWQALSHAFGRRLVLSSTFRILADLLGFAPLCIFGIYDLHGK 329
QY 354 -NDKAPADWQ-----GYFTVLLPVTAACLOTVLHQFHIQFVSGMRKIXT 397
DQ 330 ENDPQKPTQFLGVYFVSSQEFANAAVAVLFLFALLLQRTFLQASVYVATIEGINDRK 389
QY 398 AVIGAVVRKA--LVITNSARKSTVGEIVNLMSVDAORFMDLATYINMISAPLOVIAL 455
DQ 390 AIOCKITNKIMHLSHLSMGEMTNGQICHLVADTNQMLMFFFLCENLWAMPQOIIYGV 449
QY 456 YLLMLNLGSPVLAGVAVVAVLWPNVAVMANKTKTYQVAAHMKSKONRIKLMNEIINGIKVL 515
DQ 450 ILLYIILGVSALIGAAVIILLAPQVFAVATKLQAQRSTLEYSNBRUKQTNEMRGIXKL 509
QY 516 KLYAMELAFKDKVLAIRQEBKVLKKSAYLSAVGTFTWVCPTFLVALCTFAVYVYTIEN 575
DQ 510 KLYAMENIFRTRETTTRKETSIRAPAIYTSISIFMNTAIPLAAVLITFEVGHVSFEKA 569
QY 576 ILDAQTAFVSLALFNILRFPNLIMPVTSIVQASVSLKRLRIFLSHBELEPDSI----- 630
DQ 570 DFSVSAFASLSLPHILVTLPLFLLSSVRSIVKALVAVQKLSFELSSABEIREBCAPHEP 629
QY 631 -----ERRPVK-----DGGGTN--SITVRNATTTWA 654
DQ 630 TPQGPASKYQAVPLRVNRRKPAEDCRGLTGLQSLVSPSADGDACCQVQIMGQYFTWT 689
QY 655 RSDPPTNGITFRSIPEGALVAVVGVQCGKSSLLSALLAEDKREG----- 700
DQ 690 PDGIPTLSNITIRIPRQQLTWIVGVQCGKSSLLAALGEXQKVSAGVFWSSLPDSEIGE 749
QY 701 -----HVAIGSVAVVPOQAMIQDSIRENILEFCQLEBPYRSVIOACALLP 748
DQ 750 DSPERETADDDIRKGPVAVYASQKPMNLNATVEENIIPESPNNQRYKRVIVACSLQP 809
QY 749 DLEILPSGDRTEIGEKGVNLSGGQKQVSLARAVYSNADYILFDDPLSADAVAHYKHIIE 808
DQ 810 DIDILPHGDOQTQGERGINLSGGQRRISVARALYQAHANVFLDDPSALDIHLSHLMQ 869
QY 809 NVTPKKGMLK-----NKTIRLIVTHSMGYLPQVNVITYMSGKISIMKSYQSLLRDGAFA 863
DQ 870 -----AGILELRLDRDKRTVVLVTHKLQYLLPHADWITAMKQGTIQREBTLKDFQSRSEQLF 924
QY 864 EFLRTVASTBOEODAEENGVTGVSGPGEAKOMENGMVLTDSAGKOLQROLSSSSSYSGD 923
DQ 925 EHKKTILMN-RQDQLEKEITYT-----ERKATEPQG-----LSRAMSSRDGL--- 965
QY 924 ISRHNSYAELOKAKEKETWKLMEADKA-----QTGQVQLSVYWDYMAKIGLFIISLS 978
DQ 966 -----LQDBEEEBEAAESEBDNLSMLHQBAEIPWRACAKYLSAGILLISLT 1015
QY 979 IFLMCHNVSAALASNYMLSTWTDPIV-----NGTQET-----KRLSYGALGI 1024
DQ 1016 VFSQLLKRMVLAADIDYWLAKWTDSALTTPAARNCSISOECTLDQTVYVAMVFLYCSGI 1075
QY 1025 SGIIVAGVGSMAVSIIGGILASRCHVDLHSLRSFMSFFERTPSGNLVNFRSEKELD 1084
DQ 1076 ---VLCVTSVTVENTGLKVAKRHLRSILNRIILAAMPFEFTTBLGSLNRFSSDCNTID 1132
QY 1085 SMIEPVIKMFMSGLFNVIIGACIVILLATPIAIIIPPLGIYFVQRFYVASSRQKLE 1144
DQ 1133 QHISTTECLSRSTLLCVSALAVISYVTPVFLVALPLAIVCYFIQCFRVASRDLOQLD 1192
QY 1145 SVSRSPVYSHENFTLGVSVIRAFEBQERFIHQSDLKVDENQAKAYYSIVANRWLAVLE 1204
DQ 1193 DDTQPLSLSHFAETVEGLTTIRAFRYEARFOQKLEVTDSNNIASLFLTANRWLEVRME 1252

```

Qy 1205 CVGNCIYVFAALPAVIS--RHSLAGLVGLSVSYLOVTLTNMLVRMSSEMETNIVAE 1262
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1253 YIGACVLIIVAVTSTISNSLHRELSAGVLGLITLALMNSYLNWVRNLADMEIQOLGAVK 1312
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1263 RLKEYSETEKE-----APWQIOETAPSPSPVOVGFERNYCLRYREDLDFVLRHINVT 1316
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1313 RHIGLTKTEASYSGLLAPSLI-----PKMPPDQCKIOIQNLSTRYDSLSLPAVKHVAL 1367
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1317 INGEKGVIVRTGAGKSSLGLGLFRINESABEIIIDGINIAKIGLHDLRFKTTIIPD 1376
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1368 ISPOKIGICGRTSGKSSFSLIAFFRMYDTFEGHIIIDIGIRKLPLHATLPSRLSIID 1427
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1377 PVLSGSLRMNLDPFSQVDEEVMVTSLEHLKDPVSALPKLDBHECAGGSEINSVGQRQ 1436
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1428 PVLSGSLRMNLDPFSQVDEEVMVTSLEHLKDPVSALPKLDBHECAGGSEINSVGQRQ 1487
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1437 LVCLARALLRKTKLIVDEATAVDLETDLIQSTIRTOFEDCTVLTIARHLNTIMDYTR 1496
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1488 LFCALARAFVKTSTIFINDEATASIDMATEINLOKVMTAFADRVYTIARHVTILSADL 1547
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1497 VIVLDKGEIOEYGA PSDLLQOR-GLFYSMAK 1526
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1548 VIVLKRGAILEFDPDEKILSRKDSVFASFVR 1578
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 5
US-10-995-561-526
; Sequence 526, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 526
; LENGTH: 1404
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-995-561-526

Query Match 13.5%; Score 1341.5; DB 6; Length 1404;
Best Local Similarity 27.8%; Pred. No. 1.9e-90;
Matches 369; Conservative 242; Mismatches 473; Indels 243; Gaps 34;

Qy 32 CFQNTTVLVWVPCFYLMACFPYFL-YLSRHDRGYIOMTPLNKTKYALGF---LLMTVCW 86
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 26 CFVDALNVVPHVFLFTFPILFIGWGSQSSKVIHHS-----TWLHPGHNLRWILTF 79
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 87 ADLFYSFWEBSRGIFLAPV-----FLVSPF---LGITTLATFLIQLERRKGVSSG 136
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 80 MLFLVLCETABEGLISGCVESHHLLHMPAGMFAAVTSV---YHNHETSNEPKLLI 137
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 137 IMLEFWLVAVCALAIRSKIMTALKEDAQVDLFRDITF-----YVFSLLIQLVLS 189
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 138 ALLVYMLAFI-----TKIKFVKFLDAHIGSQRFCLTGLLVLLIGMLLIVE--VN 188
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 190 CFSGRSLFSETHIDNPCEBS-----SASPLSRITFWITGLVIGYRQPLRG 238
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 189 VIRVRRITFFKTPREVR-PEDLODLGVRLQPFVNILSKGTWYMNNAFIKTAHKPI-- 245
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 239 SDLSLNKEDTSEGVVFLVK---NMKKECAKTRKQPVKVYSSKDPAQPKESKVDANE 295
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 246 -DRAIRG-----LPIARALTNVQRLC-----E 268
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 296 EVELLIYKSPQ-KEMNSLFLVLYKTGYFVLMSPFKAIHDLMMFSGPOLIKLLIKFV- 353
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 269 AFDAQVKDIOGTGABAIWQALSHAFGRRLVLSSTRILLADLLGFGAGPLCFTVDHLG 328
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

Qy 354 --NDTKAPDMQ-----GYFTVLLFVTAQLOTVLHOYFHHCFVSGMRK 396
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 329 KENVPQKPTQFLGVFVSSQOEFLANAYVLALFLLLORTLQASYYVAIETGINLR 388
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 397 TAVIGAVYRKA--LVITNSARKSSTGEIVINMSVDARFMDLATYINMISAPLOVLA 454
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 389 GAIOTKYNKIMHLSNLSGEMTQAQICULVADINQMLMFFFLCPNLMAMVQIIVG 448
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 455 LYLLMLNIGPSVLGAVMVLMPVNAVMAMKTKTYQVAHMKSDNRIKLMNEILANGIV 514
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 449 VILLIYLLGVSAIIGAIVITLAEVQFVATKLSQAORSTLEYENRILKQENMLRGIKL 508
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 515 LKTYAMELAFKDKYALROEBLKYLKXSATYLSANGTCTVCTPVLALCTFAVYVTTDEN 574
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 509 LKTYAMENITRTRVETTRRREKMSLRAFAIYTSISFMNTAIPAAVLTFFVGHVSFKE 568
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 575 NIIDAQAFVSLAFNITLRPLNLTLPWISSIYQVASYLKRIRLFLSHEBLEPDSI---- 630
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 569 ADSPSVAFASLSLFIHLVPLFLSSVSSTYKALVSVQKLSFELSABIREQCAPIRE 628
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 631 -----ERRPVK-----DGGCTN-SITVNAATFTW 653
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 629 PTPQSPASKYQAVPLRVNKRKRPAREDCRGLTGPLQSLVPSADGDADNCCVQINGYFTW 688
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 654 ARSDPPTLNGITRISREGALVAAYGVQVCGCKSSLSLLALEMDEKVEG----- 700
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 689 TPDDGIPPLSNITIRIPGQOLMTIVGQVCGCKSSLSLLALEMDEKVEG----- 748
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 701 -----HVALKGSVAAYVPOQAWIIONDSLRNIIIFGCGLEBPYRSVYQACALL 747
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 749 EDPSPEREKTATDDIRKRGVAYASQCPMLNATVEENIIFESPFRKORYKMTVEACSLQ 808
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 748 PDLEILPSGDRTEIGEKVNLSSGQKQVSLAARVSNADIIYFPDDLASVADAHVKGIF 807
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 809 PDIDILPHGQOTQGEINGINLSGQQRISVAYRALYGHANVFLDDPFSALDILHLSHM 868
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 808 ENVIYKPGMKL-----NKRILVTHSMSTYLPQVDVITVMSGGKISEMSTYOELLARDCAF 862
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 869 Q-----AGILELLRDDKRTVVLVTHKLYPHADWIIMAKDGTIOREGTLKDFORSQOL 923
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 863 AEFRTYASTEOBDAEENGVTGVSQPKKAKQENGMVLTDSAGKQLOKORLSSSSYSYG 922
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 924 FEHWKTLMN--RODELEKETVT-----ERRATEPQO-----LSRAMSSRDGL-- 965
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 923 DISRHNSIAELQKAEKKEETWKLMEADXA-----QTGQVKLSVYWDYMKALGLFTSFL 977
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 966 -----LQDEEBEEDAAESEEDDNLSSMLHQRAEIPRACAKYISSAGILLLSL 1014
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 978 SIFLPMGNHVSALASNTWLSLMTDDPIV-----NGTOEHT-----KVLRSVYALG 1023
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1015 LVFSOLKJHNVLAVIDWMLAKMTDSALTLTPAARNCSSISOECTLDQTVYAMVFLVCSLG 1074
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1024 ISGIAVFGVSMAYVSIQGLASRCILVVDLHSLRSFMSFEETPPSGNINLRPSKEIDTV 1083
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1075 I---VLCVTSVVEWTEGLVAARLHRSLSNRLITLAPMRFETTPGSLINRRPSDDNTI 1131
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1084 DSMIPEVIKMFMSLNFVAGACIVILATPIAIIIPPLGLIYFVQVFRVSSRODKRL 1143
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1132 DQHIPTBLECLSTLSCVLAALAVISYTVFVPLVALLPLAIVCYFIQKRYRVASRDQOL 1191
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1144 ESVSRSPVSHENETLIGSVIPAFEBQERFIHQSDLKVDENQKAYPSIVANRWLAURL 1203
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1192 DDTTOLPLSHFAETVEGLTIRAFRYEARFOOKLEYTDSNNIASLFTLAARWMLVEVM 1251
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1204 ECVGNCI 1210
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1252 -AIIHRCM 1257
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 6
US-10-995-561-969
; Sequence 969, Application US/10995561

```

```
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CLO01559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 969
LENGTH: 410
TYPE: PR1
ORGANISM: Homo sapiens
US-10-995-561-969

Query Match      9.8%; Score 969; DB 6; Length 410;
Best Local Similarity 50.8%; Pred. No. 7,1e-64;
Matches 197; Conservative 66; Mismatches 121; Indels 4; Gaps 2;

QY 495 MSKDNRIKLNNEILNGIKVLYAMFLAFKDKVLAIRQELKVKKSAVLSAVGTPTWV 554
DB 1 MRQDSRRLTSSILRNKTKTFHMEGAFIDRVLGITGQELGALRTSGLLFVSUYSFQ 60
QY 555 CTPLVALCTPAAVYVTTIDENNILDAQTA FVSLALFNILRFPNLILPMVSIISIVQASVSLK 614
DB 61 VSTFLVALVAVHVLVAE-NANNAEKAFVTLTVINILINKAQAFLPFSIHSLVQARVSFD 119
QY 615 RLRLFLSHHELEPDSIERPKVGGGNTSITVRNATPTWASDPTPLNGITFSIPEGALV 674
DB 120 RLTVFLCLEEVDPGVSDSSGSAAGKCCITIGATFAMSOESPCLHRINILTVVQGLL 179
QY 675 AVGVGGGKSSLSALLAEMDKVEGVAIKGSVVVYVQQAIVQDSIRENILEFCQLE 734
DB 180 AVGVGVGKSSLSALLAELGSLKVBGFSIEGAIVYVQEAIVQNTSVENVCQGEIDP 239
QY 735 PYRVSIVQACALLPDLLELPSSGDRTEIEKGVNLGGQKQVRSVLARAVYSNADIVLPDP 794
DB 240 PMLRVLVACALQPDVDSPFGIHITSIGEGMNLSSGQKRLSLARAVYRKAAVYLDDP 299
QY 795 LSAVDARVKGKIFENVIGPKMKLKKTKRILVTSHSSTYPOVDVITVMGSGKISEGVSQ 854
DB 300 LAADAHVGHVFNQVIGPGLLOCTRIILVTHALHILPOADWITVLANGAIAEWGSYOE 359
QY 855 LLARDGAFAEFLRTYASTEOGDABENG 882
DB 360 LLQRKALVCLL---DQARQPDGRGEGG 384

RESULT 7
US-10-613-744-14
Sequence 14, Application US/10613744
Publication No. US20050272093A1
GENERAL INFORMATION:
APPLICANT: MacKinnon, Roderick
TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
TITLE OF INVENTION: Channel Proteins, and Uses Thereof
FILE REFERENCE: 018512-002901US
CURRENT APPLICATION NUMBER: US/10/613,744
CURRENT FILING DATE: 2003-07-03
PRIOR APPLICATION NUMBER: US/09/275,252
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: US 09/045,529
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: US 09/054,347
PRIOR FILING DATE: 1998-04-02
PRIOR APPLICATION NUMBER: WO PCT/US99/06307
PRIOR FILING DATE: 1999-03-22
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
```

```
SEQ ID NO 14
LENGTH: 407
TYPE: PR1
ORGANISM: Homo sapiens
US-10-613-744-14

Query Match      8.9%; Score 883.5; DB 6; Length 407;
Best Local Similarity 48.7%; Pred. No. 1.4e-57;
Matches 167; Conservative 74; Mismatches 95; Indels 7; Gaps 4;

QY 1551 LTRLAEDPAEPRY-----RTRRRARFVSKGNCVAKHNRQEGRFQDQVPTTVLVDKMP 1606
DB 16 LPQARDL-PRHISRRRTKRIQRYRKQCKVHGNVETRYRLTDITVLVDLKR 74
QY 1607 HTLLIFTMSPFLCSWILPAMWMLIAFAHGLDPAPEGTN-VPCTVSIHSFSSAPLFSIEVQ 1665
DB 75 FNLLIFVWVTVVTLFFGMLTWMILAYIRGMDHIEDSPWPCVTNNGVSAFLFSIETE 134
QY 1666 VTIGFGAMTSECPPLAILIIVQIVGLMINAIMLGCIFMKTQOARRAETLIFSKRAV 1725
DB 135 TTIGYRVITDKCPREGIILLIQSVLGSIVNAFMVGMFVKISQPKKRAETLVFSTYAV 194
QY 1726 ITRHGRCLCFMLRVGDLRKSMIISATIHQVVRKTSDEGRVPLHOVDIPMENGVGNG 1785
DB 195 ISMRDGKLCMFVRGDLRNSHIVASIRAKLIKSKQTSSEGFILPNOTDINVGYYTGD 254
QY 1786 IFVLAPLIIVHVIDSNSPLYDLAPSDLHHODLEIIVLBGVETTGITQARTSYLADE 1845
DB 255 LFLVSPILISHEINQOSPWFEMISKAOL-PKELEIIVLBGVETATGWTQARSSYTSE 313
QY 1846 ILMGQRFVPIVAEDGRKSYVDYSKFGNTIKVPTPLCTARQDE 1888
DB 314 ILMGYRFTPVLTLEDGFEVDYNSFHEHYETSTPSLAKELAE 356

RESULT 8
US-10-995-561-967
Sequence 967, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CLO01559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 967
LENGTH: 268
TYPE: PR1
ORGANISM: Homo sapiens
US-10-995-561-967

Query Match      8.1%; Score 799; DB 6; Length 268;
Best Local Similarity 56.3%; Pred. No. 1.2e-51;
Matches 151; Conservative 56; Mismatches 61; Indels 0; Gaps 0;

QY 1264 LKXSETEKAPWQIOETAPSSWPQGVREFFNYCLRYRDLDPVLRHINVTNGEKV 1323
DB 1 MDYAWTPKEAPWRLPCCAQPPWPOGQIEFFRGRLRPDELPLAVQGVFFKIHAGK 60
QY 1324 GIVRTAGKSSLTGLFRINESAGEIIDGINIAKIGLHDLRFKITIIPQDPLPFGS 1383
DB 61 GIVRTAGKSSLSGLRLQEAABGGIWDGVPLAHVGLTLLSRSLIIPQDPLPFGS 120
QY 1384 LRNLDPFSQYSDDEVNTSLTSLAHKDFVALPKDLHECAEGENLSVGQROLVCLARA 1443
DB 121 LRNLDDLQHSDDAIVMALETVOGLKALVSLPQOLQYKCDKREDSDSVGQKOLLCLARA 180
QY 1444 LLRTKTLIVDEATAVADLETDLLIOSTIRQPEDCTVLTIAHRLNTIMDYTRVIVLDXG 1503
```

```

Db      181 LIRKQIILLDEBATAAVDPBTELOQMAMGSMFAOCTVLLIAHRLRSVMDCAFLVYMDKG 240
QY      1504 EIQEGAPSDLLQQRGLFRYSMAKDAQLV 1531
        ::|::|::|::|::|::|::|::|
Db      241 QVAESGSPAQLLAQGLFRLDAQESLTV 268

```

RESULT 9
US-10-613-744-13

```

Sequence 13 Application US/10613744
Publication No. US20050272093A1
GENERAL INFORMATION:
APPLICANT: Mackinnon, Roderrick
APPLICANT: The Rockefeller University
TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
TITLE OF INVENTION: Channel Proteins, and Uses Thereof
FILE REFERENCE: 018512-002901US
CURRENT APPLICATION NUMBER: US/10/613,744
CURRENT FILING DATE: 2003-07-03
PRIOR APPLICATION NUMBER: US/09/275,252
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: US 09/045,529
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: US 09/054,347
PRIOR FILING DATE: 1998-04-02
PRIOR APPLICATION NUMBER: WO PCT/US99/06307
PRIOR FILING DATE: 1999-03-22
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 391
TYPE: PRN
ORGANISM: Rattus norvegicus
IS-10-613-744-13

```

Query Match	7.9%;	Score 786.5;	DB 6;	Length 391;
-------------	-------	--------------	-------	-------------

```

QY      1564  RTRERRARFVSKKCNCAVNAHKNIEQGRFL--ODVFLLVYLKMPHTLLIFMSPGCSML 1622
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      35   RSRQ-RARLVSEKGRCHNIEFGNVDAQSRFFIFVDITVTLVLDKRYKMTVFITATLGSNF 93

QY      1622  LEAVVMMLIAFAHGDL-----APGGETNPVCTS.IHSFSSAPFLSIEVOVTTIGFGGRWTE 1677
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      94   LFGLLMTVVAVVHNDLPEFYPPDNR--PCVENINGMTSAFLSFLETQVTIGYGRFPVTE 151

QY      1678  ECPALITILLYONIVGMINAIMGCI-FMKTQAQNRARETLIFSHAVIITLRHGRICML 1737
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      152  QCATRAIFELLTQSLIGVINSFPGCAILAKTSRPEKGRKTTTFSSNAVISRGGCLLI 211

QY      1738  RVGDLARKSMIISATIHMQVARKTSSPEGEVPLHQVIDIPMENGVGNGCIFVAPLIYHV 1797
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      212  RVANLFRKSLIGSHYQKLTKTITTPBGETIILIDQTNINFPVADGNEMLFISPLTIYHI 271

QY      1798  IDNSSPYDLAPSLDHHODLEIIVYLEGVVETTGITTOARTSYADEILMGQRFVIVA 1857
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      272  IDHNSPFFHMAEYTL-SQODELVFLDGTVESTATCOVTSYVBEELVLMGRYFVPIVS 330

QY      1858  E-EDGRKSYVDYSKFGNTIKVETPLPC 1881
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      331  KTKGKGRVDPFNFNGKIVETTPHC 355

RESULT 10
US-10-793-626-326
; Sequence 326, Application US/10793626
; Publication No. US20050255476a1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348005

```

RESULT 10
US-10-793-626-326
; Sequence 326, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P034800S

```

: CURRENT APPLICATION NUMBER: US/10/793,626
: CURRENT FILING DATE: 2004-03-04
: PRIOR APPLICATION NUMBER: 60/164,258
: PRIOR FILING DATE: 1999-11-09
: NUMBER OF SEQ ID NOS: 4472
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 326
: LENGTH: 548
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: synthetic
: OTHER INFORMATION: amino acid sequence
US-10-793-626-326

```

Query Match	5.2%	Score 514.5	DB 6	Length 548
Best Local Similarity	27.1%	Pred. No. 3e-30		
Matches 157	Conservative 110	Mismatches 211	Indels 101	Gaps 16

Qy	1002	DPINVGQEHKRVKRLSYVGLAQLSQGLAVFQSYMAVNSIGILASRLCHVDLLHSILRSPM	106
Db	12	DGVIN---NHSLTQOEKFSHLGVAIGALNF-----IFLIVRPPI	47
Qy	1062	SFFE---RTPEGNIIVNRFSKEL-----DTVDSMTPEVIK-----MEMG	1096
Db	48	EFITQYLAQMTSNKILVIDIRKQLVNHQLALSVRFPYANNQGVQVLSRVINDVEQTKDFILT	107
Qy	1097	SLFNIVGACIVILLATPI-----AAIIIPPLGI--IYFVQRFYVASSRQKLE	1144
Db	108	GLAMIMIDCITIIILATISMFPLDKLAFPAIFIPFYLIVYFPFGRL-----KRLTVR	162
Qy	1145	SVSSPFPYSHNELLGVSIVIRAEBOERFTHQSDLKVDENQKAYPESIVANRMLAVLE	1204
Db	163	SQLAEVQGFPLHERVQGMVSIVKSPALIEDNEAKNFD--NHNKNFLQAFQHTRNVAQSPA	219
Qy	1205	CV-----GNCIVAFALPAVISHSHSAGLVGLSVSYSLQVTTYLMVLVMSSEMETNI	1255
Db	220	AIINTVDLGPFIIVGVGSYLAIT--GSTITVGTIAAFVGLYEDLFGRLRLVSSFTTLTOSF	278
Qy	1259	VAVEREKYSETEKE-----APWQIQETAPPSWPDQVGEFRPNYCLARREDLDFTLR	1311
Db	279	ASMDRVQLMDEDEDVDINGIGAGQIKISK-----QGIDLKHSYFKNEKEKVELH	328
Qy	1312	HINVTINGEKGVIVGTGAGKSSLTGLFRINESAGEIITIDGINAKTIGLHIDRFKIT	1371
Db	329	DIMLTINKGEIVAFVMSGGSKSTLIMLIPFYDVDTQGEIILDHNVWDPLTGSILRNIG	388
Qy	1372	IIPQDPLFEGSGLRML---DPFQSYDEDEFWTSLELAHLKDPFVSALPDKLDHCEAGEE	1422
Db	389	LVOQDNLIFSPTVAKENITLLGRP--DAVDDEVEAKAKANAHDFISLNPNGYDEBGEVGV	446
Qy	1429	NLSVQROVLCLARALRKTKILVLEDEATAVADLETDLLQSTRTQPEDCVTLIARL	1488
Db	447	KLSSGGQKORISIAIFILNPPVILLIDERTSLDSESAIIIOEALDVLSSKRTLLIVARLL	506
Qy	1489	NTIMDYTRVIVLDKGEIOEYGAPSDLLQQRGLF---YSM	1524
Db	507	STIHDRIYVMENGRIVETGTHTQQLINKGAYEHLXSI	545
<p>RESULT 11</p> <p>US-10-467-657-1346</p> <p>Sequence 1346, Application US/10467657</p> <p>Publication No. US20050260581A1</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: CHIRON Spa</p> <p>APPLICANT: FONTANA Maria Rita</p> <p>APPLICANT: PIZZA Mariagrazia</p> <p>APPLICANT: MASIGNANI Vega</p> <p>APPLICANT: MONACI Elisabetta</p> <p>TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS</p> <p>FILE REFERENCE:</p> <p>CURRENT APPLICATION NUMBER: US/10/467,657</p>			

```

RESULT 11
US-10-467-657-1346
; Sequence 1346, Application US/10467657
; Publication No. US20050260581A1
GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657

```



```

; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 575
; LENGTH: 701
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-995-561-575

```

Query Match	4.2%;	Score 419.5;	DB 6;	Length 701;
Best Local Similarity	25.2%;	Pred. No. 4.3e-23;		
Matches 149;	Conservative 130;	Mismatches 234;	Indels 79;	Gaps 21;

```

QY 977 L$1L$1F$M--C$H$V$A$M$N$M$1$M$T$D$P1$V$G$T$O$E$H$K$V$L$S$V$G$A$1$G$1$Q$1$V$F$G$S 1033
Db 140 L$1L$F$L$V$L$V$L$S$1$G$E$M$A1$P$F$T$G$R$1$T$W1$I$D$D$G$A$D$T$R$U$L$1$M$S1$I$1$T$1$A$V$1$E$F$--- 196
QY 1035 M$A$V$1$G--G1$L$A$R$C$L$H$V-----D1$H$1$1$L$S$P$M$S$F$F$E$R$T$S$G$1$V$N$F$S$K$E$L$D$V$D$S$M1 1087
Db 197 ---V$G$D1$Y$N$N$T$M$G$H$V$H$1$Q$E$V$F$G$A$1$R$O$E$T$E$F$F$O$O$N$O$T$G$1$N$M$S$R$V$E$D$T$1$S$D$S$L 252
QY 1088 P$E$V1$K$M$E$G$S$F$E$N1$G$A$C1$V--1$L$-----A$P$1$A1$1$1$P$P$1$G$1$Y$F$F$O$R$Y$V$A$S$R$O$L$K 1144
Db 253 S$E$N1$S$1$F$1$W$Y1--V$R$G$1$C$L$1$G1$1$M$1$M$G$S$V$1$T$W$1$1$T$1$L$P$L$F$1$P$K$G$K$Y$O$L$1$E$V$Y$R 310
QY 1142 R$L$E$V$S$R$S$P$V$Y$H$F$E$T$1$L$G$V$1$R$A$F$E$O$E$R$F1$H$O$S$D$L$K$Y$D$E-----N$O$K---A$Y$P$S1$V 1194
Db 311 --E$1$A$K$S$S$Q$V$A--1$E$A$1$M$A$P$Y$R$S$F$A$N$E$G$E$A$K$F$R$E$K1$O$E1$K$1$M$O$K$A$V$A$Y---A 362
QY 1195 A$N$R$V-----L$A$V$L$E$C$V$N$C1$V$L$F$A$L$P$A$V1$R$H$S1$A$G$L$V$G$L$S$Y$S$1$O$V$T$Y$1$N$W 1246
Db 363 V$N$S$W$T$S1$1$G$M$L$K$V$G1$Y1$G$Q$L$Y$T$S$G$A-----V$S$G$1$N$1$V$T$Y$1$Q$M$O$F$Q$A$V$E$V 413
QY 1247 L$V$R$S$S$E$M$E$N1$V$A$V$E$R1$A$K$E$S$E$F$E$K$A$P$M$O1$O$E$A$P$S$---W$P$O$V$G$R$V$E$F$R$N$Y$C1$R$Y 1302
Db 414 L$L$1$Y$P$R$O$K$A$V$G$S$E$K1$F$E$Y$L$D$T$P$R-----C$P$S$G$U$L$P$L$H$E$G1$V$O$F$O$D$V$S$F$A$Y 465
QY 1303 --R$E$D$L$F$V$L$H$1$N$W$1$T$1$N$G$E$K$V$G1$V$G$T$G$K$S$L$T$G$L$P$R1$N$E$S$A$G$E1$1$D$1$G$1$N$F$A$K 1366
Db 466 P$N$R$D$V--L$V1$O$G$L$F$T$L$R$G$E$V$1$V$L$V$G$N$G$S$K$T$V$A$L$L$O$N$L$V$O$P$T$G$O$1$D$K$P1$P$Q 524
QY 1361 I$G$L$H$L$R$F$K$1$T1$I$B$O$D$P$V$L$F$G$S$L$R$M$N$L--D$P$S$O$Y$D$E$E$V$W$T$S$1$E$L$A$H$L$D$F$V$S$A$L$P$D$K 1418
Db 525 Y$E$H$Y$L$R$Q$V$A$V$G$B$P$Q$V$F$G$R$S1$O$E$N1$A$Y$G$L$T$O$K$P$M$E$1$T$A$A$V$S$G$A$H$S1$G$L$P$Q$ 584
QY 1419 L$D$H$C$A$R$G$E$N$L$S$V$G$O$R$O$L$V$L$A$L$A$L$R$K$T1$K1$V$L$D$E$A$T$A$V$D---L$E$T$D$L1$Q$1$S$T$R$T 1474
Db 585 Y$D$T$V$D$E$A$G$O$L$S$G$O$R$O$A$V$A$L$A$R$L$1$R$P$C$V$L1$L$D$D$A$T$S$L$D$N$S$Q1$O$V$Q$1$L$Y$E$S--P 642
QY 1475 Q$E$D$C$Y$V$L$T$1$H$R1$N$T1$M$D$Y$R$V1$V$L$D$K$E1$O$E$V$G$A$P$E$L$D1$O$O$R$G$L$P$F$S$M$A$K 1526
Db 643 E$Y$R$S$R$V$U$L1$Q$H$L$S$L$V$E$Q$A$D$H1$L$F$L$B$G$A1$R$E$G$H$O$Q1$M$E$K$G$C$Y$W$A$M$V$O 694

RESULT 14
US-10-995-561-574
Sequence 574, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: C1001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 574
LENGTH: 808
TYPE: PR1

```

ORGANISM: Homo sapiens
US-10-995-561-574

Query Match	4.2%	Score 419.5;	DB 6;	Length 808;
Best Local Similarity	25.2%	Pred. No. 5.3e-23;		
Matches 149;	Conservative 130;	Mismatches 234;	Indels 79;	Gaps 21

Qy		977	LSLFLP--CNHVSALSNWLSLMTDPIVNGOEHKTVLSYGAIGISOGIAVPFGYS	1034
Dd		247	LSLFLVLVSLSSGEMAIPEFTGRLLTWMLLODSADPTFRLLTMSILTTIASVALEF---	303
Qy		1035	MAYSIG-GILASRCLHV-----DLHSILRSPPSFERRTPSGNLVNRFSKELDTVDSMI	1087
Dd		304	----VEGDGIYNNTGHVSHSLOGEVFGAVLRQETFEFOQNQTGNMRSVTEDTSTLSDSL	359
Qy		1068	PVIYKRMGSLFNVIAGCIV-ILL-----ATPLAIIIPPLGLIYFFQRYYVASRODK	1144
Dd		360	SENISSLPLMWL--VRGCLLGIMLMGSVSLMTWLLITPLLFLLPKKGXWYOLLEVVOR	417
Qy		1142	RLESVSRSPVYASHNETILGVSIVRAFEEOERFIHQSDLKVDZ---NOK---AYPSIV	1194
Dd		418	--ESLASQVA--IEHLSAMPTVRSFANEGBEQKRKEQLIKTLNQKAVALY-----A	469
Qy		1195	ANRW-----LAVRLCYGCNCIVLPALPAVISRHSLSAGLVGSYSYLQVTTYIANM	1246
Dd		470	VNSWTTSISGMCLKVGILYIGGQLVTSGA-----VSSGILVTFLVYOMOFQAWEV	520
Qy		1247	LVMSSMENTINIVAVEBLKEYSEDEKAPMOIQETAPSS---WPQVGRPEPNYCIRY	1304
Dd		521	LSLTYPNVQAVSGSEKIPEYLDTPP-----CPSSGLTPLHLHGLVQFDVSVFAY	572
Qy		1303	--REDLDFVLRIHINVTINGEKGKGVIGRTGAKSSLTGLFPRINESAGEIITIGINIYAK	1366
Dd		573	PNRPDV-LVLOGLFTLRPEGVVALVPNGSGKSTVAALLQNLVOPRGQGLLDGKPLRPQ	631
Qy		1361	IGHDLRFKTIITIPDPVLFSSGISRMWL--DPFSQYSDERYWTSLELAHLKDQFVSALEPK	1418
Dd		632	YEHRYLRYOVAVAVGQEPQVFRSIOENIAYGLYOKPMTEBETAAVAXGANSFISGLPQG	691
Qy		1419	LDHECAEGGENLSVGQROLVCARALLRKTKILVLDATAVD---LETDDLIOSTIRT	1474
Dd		692	YDFEVNDAGSQSLGGCGRQVALMARALKRECVLLIDDTALSANDNSQLQEQLYES--P	749
Qy		1475	QFEDCTVLTIAHRLANTIMDYTRVIVLDKEGIEGYABSDLLQQGLFYSMAK	1526
Dd		750	ERSRSRVLLITQHLSLVEQADHILFLEGALIREGTTHQMLEKKGCYAMAWQ	801
<hr/>				
RESULT 15				
US-11-105-268-53				
; Sequence 53, Application US/11105268				
; Publication No. US20050260204A1				
; GENERAL INFORMATION:				
; APPLICANT: Alliant, Christian				
; TITLE OF INVENTION: ANTI-ILI-9 ANTIBODY FORMULATIONS AND USES THEREOF				
; FILE REFERENCE: 10271-126-999				
; CURRENT APPLICATION NUMBER: US/11/105, 268				
; PRIOR FILING DATE: 2005-04-12				
; PRIOR APPLICATION NUMBER: 60/561, 845				
; NUMBER OF SEQ ID NOS: 60				
; SOFTWARE: PatentIn version 3.2				
; SEQ ID NO 53				
; LENGTH: 808				
; TYPE: PRT				
; ORGANISM: Homo sapiens				
US-11-105-268-53				
<hr/>				
Query Match				
Best Local Similarity 25.2%; Pid. No. 5.3e-23;				
Matches 149; Conservative 130; Mismatches 234; Indels 79; Gaps 21;				
977 LSLFLP--CNHVSALASNWLSLMTDDPIVNGTOEHKTVLSYGAIGISOGIAVPGYS				
1034				

977 LSIFLEM--CNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYGALGISQGIAVFGYS 1034

Db 247 LSLFLVAVLSSLOEMAIPEFTGRITDMLQDGSADPTFRNLITLMSILTIASAVLEF--- 303
QY 1035 MAVSIG-GILASRCIAH-----DLHSLRS PMSFPERPSGNLVNRFSEKELDTVDSMI 1067
Db 304 ----VGDGITYNNMGHSHLOGEVFGAVLRQETFEFQONOTGNMSKRYEDTSTLSDSL 359
QY 1088 PEVIKMFMSLFENVIGACIV-ILF-----ATPIAAIIIPPLGLYFVQRYFVASSRQLK 1141
Db 360 SENISLFLMYL--VRGLCLIGIMLWGSVSLTMVTLITLPLFLPKVKGWYQLLEVOVR 417
QY 1142 RLEVSRSRSPVSHNETLGSVIRAFEBQRFIHQSDLKYDE---NOK--AYPSIV 1194
Db 418 --ESLAKSSQYA--TEALSAMPYVASFANEBGEAKFREKLOEIKTLNQKEAVAY---A 469
QY 1195 ANRW-----LAVRLECVGNCIYLFALPAVISRHSLASGLVGLSVSYSLQVTTYLNM 1246
Db 470 VNSWTTISGMLLKVGLIYIGQLVTSGA-----VSSGNLVTFVLYQMQFTQAVEV 520
QY 1247 LVNRSSEMETNIVAVERLKEYSETEKEAPWQIOETAPPS---WPQYGRVEFRNYCLRY 1302
Db 521 LLSIYPRVQKAVGSSEKIFEYLDRTPR-----CPSGGLTPLHLHGLVQFQDVSPAY 572
QY 1303 --REDLDPVLRHINVTINGSEKVGIVGRTGAKSLSLTGLPRINESAGEIITIDGINIAK 1360
Db 573 PNRPDV-LVLOGLFTLRPFGEVTLAVGPNGSGKSTVALLQNLVQPTGQLLDGKPLPQ 631
QY 1361 IGLHDLRFKTIIPQDPVLFSGSLRMLN--DPFSQYSDDEVWTSLELAHLKDFFVSLPDK 1418
Db 632 YEHRYLHRQVAAVGOEPQVFGRSLOENIAYGLTQKPTMEITAAAVKSAHSFTSGLPQ 691
QY 1419 LDHECAGEGNIYVGOQOLVCIAARALLRKTKILVLDENTAAMD---LETDDLIOSTIRT 1474
Db 692 YDTEVDEAGSQLSGGOQAVALARALIRKPCVLIIDDATSALDANSQLOVEQLYES--P 749
QY 1475 QFEDCTVLTIAHRIANTIMDYTRVYVLDKGIQEIYGAPSDLLQQRGLFTYSMAK 1526
Db 750 ERYRSRVLLITQHLSLVEQADHILFLEGAIREGGTHOOLMEKKGCVWAMVQ 801

Search completed: December 15, 2005, 15:50:05
Job time : 18.2651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:27:22 ; Search time 176.366 Seconds
(without alignments)
4800.712 Million cell updates/sec

Title: US-10-665-283-1
Perfect score: 9903
Sequence: 1 MALRGFCADSDPLMDMNV.....RSVAVAKXKPFISIPDLSL 1927

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_21:*
- 1: geneseqp1980s:*
 - 2: geneseqp1990s:*
 - 3: geneseqp2000s:*
 - 4: geneseqp2001s:*
 - 5: geneseqp2002s:*
 - 6: geneseqp2003as:*
 - 7: geneseqp2003bs:*
 - 8: geneseqp2004s:*
 - 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9903	100.0	1927	9	ADY86938 Human MRP
2	9891	99.9	1927	9	ADY86938 Human MRP
3	9873	99.7	1947	9	ADY86945 Human MRP
4	9734	98.3	1891	9	ADY86941 Human MRP
5	7860	79.4	1531	2	AAW57486 Human MRP
6	7860	79.4	1531	2	AAW74471 Human mul
7	7860	79.4	1531	3	AAW98994 Human mul
8	7860	79.4	1531	3	AAW55799 Human mul
9	7860	79.4	1531	3	AAW78873 Human mul
10	7860	79.4	1531	5	ABG61810 Multidrug
11	7860	79.4	1531	6	ABM35012 Cancer ba
12	7860	79.4	1531	6	ADB20865 MRP1 base
13	7860	79.4	1531	7	ADB87954 Human UGT
14	7860	79.4	1531	7	ADB86937 Human MDR
15	7860	79.4	1531	7	ADB92128 Human MDR
16	7860	79.4	1531	7	ADB92128 Human MDR
17	7860	79.4	1531	8	ADN97111 MRP1 prot
18	7860	79.4	1531	8	ADN97111 MRP1 prot
19	7860	79.4	1531	8	ADN97111 MRP1 prot
20	7856	79.3	1530	2	ADN99929 Human BAC
21	7849	79.3	1531	2	AAW54928 Multidrug
22	7849	79.3	1531	2	AAW54928 Multidrug
23	7849	79.3	1531	2	AAW54928 Multidrug
24	7849	79.3	1531	2	AAW54928 Multidrug

25	7849	79.3	1531	2	AAW98993 Human mul
26	7849	79.3	1531	3	AAW55798 Human mul
27	7849	79.3	1531	3	AAW78872 Human mul
28	7849	79.3	1531	3	AAW30582 Human mul
29	7763	78.4	1515	8	ADU24094 Human mul
30	7610	76.8	1489	2	AAW96952 Human mul
31	7587.5	76.6	1482	8	ADN03902 Antipor
32	7587.5	76.6	1482	8	ADN03902 Antipor
33	7572	76.5	1481	5	ADP52108 PRO poly
34	7560.5	76.3	1482	2	AAW96953 Homo sapi
35	7535	76.1	1475	9	ADX06206 Multi-dru
36	7510.5	75.8	1472	9	ADX06204 Cyc
37	7500.5	75.7	1472	2	AAW96954 Multi-dru
38	7230.5	73.0	1530	7	ADP56611 Bovine NR
39	7195.5	72.7	1416	9	ADX06208 Cyc
40	7194	72.6	1417	2	AAW96955 Multi-dru
41	7037	71.1	1368	7	ADJ69912 Human hea
42	7002.5	70.7	1528	2	AAW57487 Murine mu
43	7002.5	70.7	1528	2	AAW74472 Mouse mul
44	7002.5	70.7	1528	2	AAW98995 Mouse mul
45	7002.5	70.7	1528	3	AAW55800 Murine mu

ALIGNMENTS

RESULT 1	
ADY86938	
ID	ADY86938 standard; protein; 1927 AA.
XX	
AC	ADY86938;
XX	
DT	02-JUN-2005 (first entry)
XX	
DE	Human MRP1-mouse Klr6.2 fusion protein, SEQ ID NO: 1.
XX	
KM	Ionophore; biosensor; drug screening; diagnostic;
KW	microorganism detection; potassium channel; fusion protein;
KX	multidrug resistance protein 1; MRP1; Klr6.2.
XX	
OS	Homo sapiens.
OS	Mus musculus.
OS	Chimeric.
XX	
FH	Key
FT	Region
FT	Location/Qualifiers
FT	1..1531
FT	/note="Multidrug resistance protein 1 (MRP1)"
FT	1532..1537
FT	/note="Hexaglycine spacer"
FT	1538..1927
FT	/note="Klr6.2 protein"
XX	
XX	US2005063989-A1.
XX	
XX	24-MAR-2005.
XX	
XX	22-SEP-2003; 2003US-00665283.
XX	
XX	22-SEP-2003; 2003US-00665283.
XX	
XX	(COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX	
XX	Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
XX	WPI; 2005-252611/26.
XX	
XX	New ion channel hybrid protein, used as electrical sensor for screening
XX	an agonist/antagonist of a membrane protein and for detecting a
XX	contaminant/pollutant in a sample.
XX	
XX	Claim 25; SEQ ID NO 1; 78pp; English.
XX	
CC	The present invention relates to a hybrid protein consisting essentially

CC of the fusion of a membrane protein with an ion channel which is not
 CC naturally coupled to the membrane protein. The hybrid protein is used as
 CC an electrical sensor for screening of an agonist/antagonist of a membrane
 CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
 CC invention is also useful for medical diagnostics and microanalysis
 CC detection. The present sequence is a fusion protein comprising human
 CC multigene resistance protein 1 (MRP1) and mouse inner rectifier potassium
 CC channel Kir6.2 protein.

xx Sequence 1927 AA:

Query Match 100.0%; Score 9903; DB 9; Length 1927;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MALRGFCGADSDPLMDNNTMTNSPDTKCFQNTVAVWPCEFLTACPPFFLYLSRH 60
 1 MALRGFCGADSDPLMDNNTMTNSPDTKCFQNTVAVWPCEFLTACPPFFLYLSRH 60
 61 DRGYIOMTPLNKTKTALGFLIMVCMADLFYSFWERSRGIFLAPVFLVSPTLGITLLA 120
 61 DRGYIOMTPLNKTKTALGFLIMVCMADLFYSFWERSRGIFLAPVFLVSPTLGITLLA 120
 121 TFLIQLERRKGVQSSGIMLTFLVALVCAALIRSKINTALKEDAVDLFRDITFYVFS 180
 121 TFLIQLERRKGVQSSGIMLTFLVALVCAALIRSKINTALKEDAVDLFRDITFYVFS 180
 181 LLLIOLVSCFSDSPLESETIHDNPPCESASFLSITFMWIGLIVRGYROPLEGSD 240
 181 LLLIOLVSCFSDSPLESETIHDNPPCESASFLSITFMWIGLIVRGYROPLEGSD 240
 181 LLLIOLVSCFSDSPLESETIHDNPPCESASFLSITFMWIGLIVRGYROPLEGSD 240
 241 LMSLNKEDTSEOVYPVLVKNMKKECAKTRKQPVKVYVSSKDPAPQKSSKYDANEVAL 300
 241 LMSLNKEDTSEOVYPVLVKNMKKECAKTRKQPVKVYVSSKDPAPQKSSKYDANEVAL 300
 301 IVKSPQKEMNSLPFVYKTRGPFYLMSPFFKAHDLMMFSGPOLIKLIKRVNDTKAPD 360
 301 IVKSPQKEMNSLPFVYKTRGPFYLMSPFFKAHDLMMFSGPOLIKLIKRVNDTKAPD 360
 361 WQGFYTYLLEFYTACLOTLMVHOYPHICFVSGMRIKTAVIGAVRKALVITNSARKSSTV 420
 361 WQGFYTYLLEFYTACLOTLMVHOYPHICFVSGMRIKTAVIGAVRKALVITNSARKSSTV 420
 421 GEIYVNLVSADQRFMDLATYINMISAPLOVILALYLLMLNGPSVLGAVAVMLVMEVN 480
 421 GEIYVNLVSADQRFMDLATYINMISAPLOVILALYLLMLNGPSVLGAVAVMLVMEVN 480
 481 AVAMAKTKTYOVAAHKKSKDNRIKLMNEILNGIKVILYAMELAFQKYLAIROBELKYLK 540
 481 AVAMAKTKTYOVAAHKKSKDNRIKLMNEILNGIKVILYAMELAFQKYLAIROBELKYLK 540
 541 KSAVLSAAGTFTWCTPFLVALCTFAVYVTTDENNILLDAQTFVSLAFNIIIRPLNTLP 600
 541 KSAVLSAAGTFTWCTPFLVALCTFAVYVTTDENNILLDAQTFVSLAFNIIIRPLNTLP 600
 601 MVISSIVQASVSLKELRIFLSHEELPDSIERRPVKGGGTINSITVRNATFTWASDPT 660
 601 MVISSIVQASVSLKELRIFLSHEELPDSIERRPVKGGGTINSITVRNATFTWASDPT 660
 661 LNTGTFSPREGALVANVGQVCGKSSLSALIAENDKVEGHVAKGSVAVYVQQAIND 720
 661 LNTGTFSPREGALVANVGQVCGKSSLSALIAENDKVEGHVAKGSVAVYVQQAIND 720
 721 SLRENIILGCOLLEERYRSVIOACALLPDLLELPSGDTEIGEKVNISGGCKOKVSLAR 780
 721 SLRENIILGCOLLEERYRSVIOACALLPDLLELPSGDTEIGEKVNISGGCKOKVSLAR 780
 781 AVYSNADITLFDPLSAVDAAHVKAI FENVIGPKMKLNKTRILTVTHSKSYLPQVDVIV 840
 781 AVYSNADITLFDPLSAVDAAHVKAI FENVIGPKMKLNKTRILTVTHSKSYLPQVDVIV 840
 841 MSGKISMGSYOELLARDGAFELRTYASTEOBDAEENGVTGVSQKKAQKEMNGM 900
 841 MSGKISMGSYOELLARDGAFELRTYASTEOBDAEENGVTGVSQKKAQKEMNGM 900

841 MSGKISMGSYOELLARDGAFELRTYASTEOBDAEENGVTGVSQKKAQKEMNGM 900
 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAELOKAKEETWKLMEADKAQTGYKL 960
 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAELOKAKEETWKLMEADKAQTGYKL 960
 961 SVYWDYKAIQGLFSPISITLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSYG 1020
 961 SVYWDYKAIQGLFSPISITLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSYG 1020
 1021 ALGISOGIAVGVSAVSIIGSIIASRCLHVDLHSIIRSPMSFEPTPSGNLVNRSFSEL 1080
 1021 ALGISOGIAVGVSAVSIIGSIIASRCLHVDLHSIIRSPMSFEPTPSGNLVNRSFSEL 1080
 1081 DTVDMSIPEVIMKMGSLFNVIGACIYILLATPIAIIIPPLGIYFVGQFYVAASSROL 1140
 1081 DTVDMSIPEVIMKMGSLFNVIGACIYILLATPIAIIIPPLGIYFVGQFYVAASSROL 1140
 1141 KRLSVSRSPVYSHPNETLIGSVYIRAPFEOERTIHOSDLKVDENOKAYYPSIYANRWLA 1200
 1141 KRLSVSRSPVYSHPNETLIGSVYIRAPFEOERTIHOSDLKVDENOKAYYPSIYANRWLA 1200
 1201 VRLSCVGNCTVLFALFAVISRHSLSAGLVGLSVSYSLQVTTYLMLVNRSSSEMETNIVA 1260
 1201 VRLSCVGNCTVLFALFAVISRHSLSAGLVGLSVSYSLQVTTYLMLVNRSSSEMETNIVA 1260
 1261 VERLKEYSETEKEAPWQIOETAPSPSWPOVGRVEFRNYCJLYREDLPEVLRIHINTINGG 1320
 1261 VERLKEYSETEKEAPWQIOETAPSPSWPOVGRVEFRNYCJLYREDLPEVLRIHINTINGG 1320
 1321 EKVGI VERTGAKSSLTGLFRINESABGEIIDGINIAKGLHDLRKTIITIQDPVLF 1380
 1321 EKVGI VERTGAKSSLTGLFRINESABGEIIDGINIAKGLHDLRKTIITIQDPVLF 1380
 1381 SGLSRMNLDPFSQYSDSEWVTSLELAHKDVSALPDLDHCEKAGEGNSVGOLOVCL 1440
 1381 SGLSRMNLDPFSQYSDSEWVTSLELAHKDVSALPDLDHCEKAGEGNSVGOLOVCL 1440
 1441 ARALLRKTKILVDEATPAVDELTDLLIOSTIRFOBEDCTVLTIAHRLNTIMDYTRVIL 1500
 1441 ARALLRKTKILVDEATPAVDELTDLLIOSTIRFOBEDCTVLTIAHRLNTIMDYTRVIL 1500
 1501 DKGEIOEYGAQSDLLQORGLFYSAKQAGLVGGGGGMLSRKGIIPERYVLTILAEDPAE 1560
 1501 DKGEIOEYGAQSDLLQORGLFYSAKQAGLVGGGGGMLSRKGIIPERYVLTILAEDPAE 1560
 1561 PRYTRERARFVSKKGCNVAAHKNIRBQGRFLQDVFTTVLDLKPHTLLIFTMSPLCSW 1620
 1561 PRYTRERARFVSKKGCNVAAHKNIRBQGRFLQDVFTTVLDLKPHTLLIFTMSPLCSW 1620
 1621 LIFAMVWMLIFALHGDLPAGEGTNVPCTVSIHSESSAFLFSIEVQVITIGFGRMVTEBCP 1680
 1621 LIFAMVWMLIFALHGDLPAGEGTNVPCTVSIHSESSAFLFSIEVQVITIGFGRMVTEBCP 1680
 1681 LAIILIVONIVGLMINAIMLGIIPKTAQARPAETLIPSKHAVITLRHRLCFMLRVG 1740
 1681 LAIILIVONIVGLMINAIMLGIIPKTAQARPAETLIPSKHAVITLRHRLCFMLRVG 1740
 1741 DLKSMISATITHQVARKTSPGEVWPLHOUNIPMENGUGNGCIPVAPLIIYHVIDS 1800
 1741 DLKSMISATITHQVARKTSPGEVWPLHOUNIPMENGUGNGCIPVAPLIIYHVIDS 1800
 1801 NSPIYDLAPSDLHHODLEIIVIEGVEYETGIIITQARTSYLADIELMGQRFVPIVAED 1860
 1801 NSPIYDLAPSDLHHODLEIIVIEGVEYETGIIITQARTSYLADIELMGQRFVPIVAED 1860
 1861 GRYSVDYSKFGNTIKVPPLCTARQLDEDRSLDALTLASSRGPRLKRSVAVAKKPRS 1920
 1861 GRYSVDYSKFGNTIKVPPLCTARQLDEDRSLDALTLASSRGPRLKRSVAVAKKPRS 1920
 1921 ISPDLS 1927
 1921 ISPDLS 1927

RESULT 2
ADY86943
ID ADY86943 standard; protein, 1927 AA.
XX
AC ADY86943;
XX
DT 02-JUN-2005 (first entry)
XX
DE MRP1-mouse Klr6.2 mutant (KR370AA) fusion protein, SEQ ID NO: 6.
XX
KW Ionophore; biosensor; drug screening; diagnostic;
KW microorganism detection; potassium channel; fusion protein;
KW multidrug resistance protein 1; MRP1; Klr6.2; mutein.
XX
OS Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX
FH Key
FH Region 1. .1531
FT /note= "Multidrug resistance protein 1 (MRP1)"
FT Region 1532. .1537
FT /note= "Hexaglycine spacer"
FT Region 1538. .1927
FT /note= "Klr6.2 mutant protein"
FT Misc-difference 1907
FT /note= "Wild-type Lys substituted by Ala"
FT Misc-difference 1908
FT /note= "Wild-type Arg substituted by Ala"
XX
PN US2005063989-A1.
XX
XX 24-MAR-2005.
XX
XX 22-SEP-2003; 2003US-00665283.
XX
XX 22-SEP-2003; 2003US-00665283.
XX
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX
PI Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
XX WPI; 2005-252611/26.
XX
DR
XX
PT New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
PS Claim 25; SEQ ID NO 6; 78pp; English.
XX
XX The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Klr6.2 mutant (KR370AA) protein.
XX
SQ Sequence 1927 AA;
Query Match 99.9%; Score 9891; DB 9; Length 1927;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

61 DRGYIQMTPLNKTKTALGFLIMIVCWADLFYSFWERSRGIFLAVFLVSPTLGTTLLA 120
121 TFLIQLERRKGVSSGIMLTFWLVAVCALAIRSKTMTALKEDAQVDLPDITFYVFS 180
121 TFLIQLERRKGVSSGIMLTFWLVAVCALAIRSKTMTALKEDAQVDLPDITFYVFS 180
181 LLLIQVLVSCSDSPLESETIHDNCPRESSASFLRITFWMTGLIVRYRQPLBESD 240
181 LLLIQVLVSCSDSPLESETIHDNCPRESSASFLRITFWMTGLIVRYRQPLBESD 240
241 LMSLNKEDTSEQVVPVLVKWKKECAKTRKQPVYVYSSKDPAPKSSSKDAEVEVAL 300
241 LMSLNKEDTSEQVVPVLVKWKKECAKTRKQPVYVYSSKDPAPKSSSKDAEVEVAL 300
301 IVKSPQKEMNDSPLEFKVYKTFGPFYFLMSFFPKA1HDLMSFGPOLIKLIFVNDTKAPD 360
301 IVKSPQKEMNDSPLEFKVYKTFGPFYFLMSFFPKA1HDLMSFGPOLIKLIFVNDTKAPD 360
361 WQGYFTVTLAFTVACLOTLVLAHQVPHI CPVSGMRKTAIVGAVYRKALVITNSARKSTV 420
361 WQGYFTVTLAFTVACLOTLVLAHQVPHI CPVSGMRKTAIVGAVYRKALVITNSARKSTV 420
421 GEIVNLMSVDAQRFMDLATTYINMWSAPLOVITLALYLLMNLGSPVLAGAVMTLAMPVN 480
421 GEIVNLMSVDAQRFMDLATTYINMWSAPLOVITLALYLLMNLGSPVLAGAVMTLAMPVN 480
481 AVMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLYAMELAKDVLAIROBELKVLK 540
481 AVMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLYAMELAKDVLAIROBELKVLK 540
541 KSAVLSAVGFTTWCCTPFLVALCTFAYVYITDENNII DAOFAFSLAFNLRPLNLP 600
541 KSAVLSAVGFTTWCCTPFLVALCTFAYVYITDENNII DAOFAFSLAFNLRPLNLP 600
601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKQGGGNTTNNATFTMARSDPT 660
601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKQGGGNTTNNATFTMARSDPT 660
661 LMGITFSPREGALVAVVGVGCGKSSLSALLAEMDKVEGVAIKGSVAVVPOQAMTQND 720
661 LMGITFSPREGALVAVVGVGCGKSSLSALLAEMDKVEGVAIKGSVAVVPOQAMTQND 720
721 SLRENIIFGQLEBPYRVSIVQACALLPDLRILPDSGRTEIGEGVNLSSGQKRVSLAR 780
721 SLRENIIFGQLEBPYRVSIVQACALLPDLRILPDSGRTEIGEGVNLSSGQKRVSLAR 780
781 AVYSNADIVLPDPLSAVDHVGKHIPENYIGPKMLKNKTRILVTHSMSTLPQVDVITV 840
781 AVYSNADIVLPDPLSAVDHVGKHIPENYIGPKMLKNKTRILVTHSMSTLPQVDVITV 840
841 MSGGKISEMSYQELRLARDGAFELRTVASTBOEQAENGVTGVSQPKGAQOMENGM 900
841 MSGGKISEMSYQELRLARDGAFELRTVASTBOEQAENGVTGVSQPKGAQOMENGM 900
901 LVYDSACKQORQLSSSSSYSGDISSRHNSTAEIOKAEKKEFWKLMADKATQGYKL 960
901 LVYDSACKQORQLSSSSSYSGDISSRHNSTAEIOKAEKKEFWKLMADKATQGYKL 960
961 SVYWDVYKALGLFSIFSLIFLMCNHVSALASNYWLSMTDDPVLNTOGTHTKVRLSVYG 1020
961 SVYWDVYKALGLFSIFSLIFLMCNHVSALASNYWLSMTDDPVLNTOGTHTKVRLSVYG 1020
1021 ALGISQGIAGVGYMAVSGIILASRCLAVDLHSILRSPMSFFERTPSGTLVNRFSKEL 1080
1021 ALGISQGIAGVGYMAVSGIILASRCLAVDLHSILRSPMSFFERTPSGTLVNRFSKEL 1080
1081 DTYDSMTPEVYIKNMGSLFNVIACIVILATPAAIIPILGITYFFVORFYVASSROL 1140
1081 DTYDSMTPEVYIKNMGSLFNVIACIVILATPAAIIPILGITYFFVORFYVASSROL 1140
1141 KRLSEVSRSPVYSHFNETLIGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIYANFWLA 1200
1141 KRLSEVSRSPVYSHFNETLIGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIYANFWLA 1200

Db 1141 KRLSEVSRSPVYSHFNETTLGVSVIRAFEEQERFLHSDLKVDEKQKAYYSIVANRMLA 1200
 QY 1201 VRLCEVGNCTVLFAALFAVISRHSLSAGLVGLSVYSIQVTTYTNMLVRLMSSEMETNIVA 1260
 Db 1201 VRLCEVGNCTVLFAALFAVISRHSLSAGLVGLSVYSIQVTTYTNMLVRLMSSEMETNIVA 1260
 QY 1261 VERLKEYSERTEKAPWQIQETAPPSWQVGRVFEFRNYCLARREDLDVLRHINVTINGG 1320
 Db 1261 VERLKEYSERTEKAPWQIQETAPPSWQVGRVFEFRNYCLARREDLDVLRHINVTINGG 1320
 QY 1321 EKVGIIVGTGAGKSLTLGLFRINSAAGEIIIDGINAKIGLHDLARKITIIIPDDPLF 1380
 Db 1321 EKVGIIVGTGAGKSLTLGLFRINSAAGEIIIDGINAKIGLHDLARKITIIIPDDPLF 1380
 QY 1381 SGLSRMNLDPFSQYSDSEEWMTSLBLAHLKDFVSALPDLHDECAEGENLSVGORQLVCL 1440
 Db 1381 SGLSRMNLDPFSQYSDSEEWMTSLBLAHLKDFVSALPDLHDECAEGENLSVGORQLVCL 1440
 QY 1441 ARALIRKTKIIVDEATNAVLETTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRIVVL 1500
 Db 1441 ARALIRKTKIIVDEATNAVLETTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRIVVL 1500
 QY 1501 DKGEIOEYGAESDILQORGLFYSMAKADGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
 Db 1501 DKGEIOEYGAESDILQORGLFYSMAKADGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
 QY 1561 PRYTRERARFVSKKGCNVAAHKNIROGRFLQDVFTTLVDLKKPHTLIIPTMSFLCSW 1620
 Db 1561 PRYTRERARFVSKKGCNVAAHKNIROGRFLQDVFTTLVDLKKPHTLIIPTMSFLCSW 1620
 QY 1621 LLFAMVWMLIAFAHGDLAGEGTNVPCVTSIHSSASALFSEVQVITGFGGRWTECEP 1680
 Db 1621 LLFAMVWMLIAFAHGDLAGEGTNVPCVTSIHSSASALFSEVQVITGFGGRWTECEP 1680
 QY 1681 LAILIIVQNVGLMINAIMGCIEMKTAQARARETIIIFSRAVITTLRHGLCFMLAVG 1740
 Db 1681 LAILIIVQNVGLMINAIMGCIEMKTAQARARETIIIFSRAVITTLRHGLCFMLAVG 1740
 QY 1741 DLKRSMTISATIHMOVAKTTSPEGEVPLHQVDIPMENGVGNGCIPVAPLIIVHIDS 1800
 Db 1741 DLKRSMTISATIHMOVAKTTSPEGEVPLHQVDIPMENGVGNGCIPVAPLIIVHIDS 1800
 QY 1801 NSPLVDLAPSLHHQDLEIIVILEGVETTGITTOARTSYADELLMGORFVPIVAED 1860
 Db 1801 NSPLVDLAPSLHHQDLEIIVILEGVETTGITTOARTSYADELLMGORFVPIVAED 1860
 QY 1861 GRYSVDYKPGNTIKVPTPLCTAROLDDEDRSLDLATLASSRGPRLKXSVANAKKPKFS 1920
 Db 1861 GRYSVDYKPGNTIKVPTPLCTAROLDDEDRSLDLATLASSRGPRLKXSVANAKKPKFS 1920
 QY 1921 ISPDLS 1927
 Db 1921 ISPDLS 1927
 RESULT 3
 ID ADY86945 standard; protein; 1947 AA.
 XX
 AC ADY86945;
 XX
 DT 02-JUN-2005 (first entry)
 DE Human MRP1-mouse Kir6.2-HA fusion protein, SEQ ID NO: 8.
 XX
 KW Ionophore; biosensor; drug screening; diagnostic;
 KW microorganism detection; potassium channel; fusion protein;
 KW multidrug resistance protein 1; MRP1; Kir6.2.
 XX
 OS Homo sapiens.
 OS Mus musculus.
 OS Chimeric.
 OS Unidentified.

XX Key
 FH Location/Qualifiers
 FT 1. .1531
 FT /note= "Multidrug resistance protein 1 (MRP1)"
 FT 1532. .1537
 FT /note= "Hexaglycine spacer"
 FT 1538. .1947
 FT /note= "Kir6.2 protein containing HA epitope peptide"
 FT 1651. .1659
 FT /note= "HA epitope peptide"
 PN US2005063989-A1.
 PD 24-MAR-2005.
 PF 22-SEP-2003; 2003US-00665283.
 PR 22-SEP-2003; 2003US-00665283.
 PR (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 PA Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
 PI WPI; 2005-252611/26.
 PS New ion channel hybrid protein, used as electrical sensor for screening
 PT an agonist/antagonist of a membrane protein and for detecting a
 PT contaminant/pollutant in a sample.
 XX Claim 25; SEQ ID NO 8; 78pp; English.
 CC The present invention relates to a hybrid protein consisting essentially
 CC of the fusion of a membrane protein with an ion channel which is not
 CC naturally coupled to the membrane protein. The hybrid protein is used
 CC as an electrical sensor for screening of an agonist/antagonist of a membrane
 CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
 CC invention is also useful for medical diagnostics and microorganism
 CC detection. The present sequence is a fusion protein comprising human
 CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
 CC channel Kir6.2 protein containing HA epitope peptide.
 XX
 SQ Sequence 1947 AA;
 Query Match 99.7%; Score 9873; DB 9; Length 1947;
 Best local similarity 99.0%; Pred. No. 0;
 Matches 1927; Conservancy 0; Mismatches 0; Indels 20; Gaps 2;
 QY 1 MALRGFCADSDSDPLMDNNTWNTSNPDFTFCFQNTVLVWVPCFYLACFPFFLYLSRH 60
 Db 1 MALRGFCADSDSDPLMDNNTWNTSNPDFTFCFQNTVLVWVPCFYLACFPFFLYLSRH 60
 QY 61 DRGYIOMTPLNKTALGFLIMVCMADLFYSFERSRGIFLAPVFLVSPILGITTLIA 120
 Db 61 DRGYIOMTPLNKTALGFLIMVCMADLFYSFERSRGIFLAPVFLVSPILGITTLIA 120
 QY 121 TFLIQLEBRKGVQSSGIMLTFWLVALVCAILRSKIMTALKEDAQVDLFRDITFYVYS 180
 Db 121 TFLIQLEBRKGVQSSGIMLTFWLVALVCAILRSKIMTALKEDAQVDLFRDITFYVYS 180
 QY 181 LLLIQVLVSCFSDSPFLFSETIHDNPPCESASGLSITTFWMTGLIVRGROQLBESD 240
 Db 181 LLLIQVLVSCFSDSPFLFSETIHDNPPCESASGLSITTFWMTGLIVRGROQLBESD 240
 QY 241 LMSLNKEDTSEQVVPVLAKMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
 Db 241 LMSLNKEDTSEQVVPVLAKMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
 QY 301 IVKSPQKEMNSLFLVLYKTEGPFYFLMSFFKAIHDLMMFGSPQILKLLIKFVNDTKAPD 360
 Db 301 IVKSPQKEMNSLFLVLYKTEGPFYFLMSFFKAIHDLMMFGSPQILKLLIKFVNDTKAPD 360
 QY 361 WQGYFTYTLFVYTAQLQTLVLAHQYHICFVSGMKIKTAVIGAVRKALVITNSAKSSTV 420

Db 361 WOGFYTVLLFVTAQLOTVLHOFHICFVSGMKRIKTAIVAGVYRKALVITNSARKSTV 420
 Qy 421 GEIYNLMSVPAORFMDLATYINMIMSAplovtlalytlmnlGSGVLAVGVAMTAMPVN 480
 Db 421 GEIYNLMSVPAORFMDLATYINMIMSAplovtlalytlmnlGSGVLAVGVAMTAMPVN 480
 Qy 481 AVMAKRTYOVAMHKSNDRIKIMNEILNGIKYIKLYAMELAFKDKVLAIROBELKYLK 540
 Db 481 AVMAKRTYOVAMHKSNDRIKIMNEILNGIKYIKLYAMELAFKDKVLAIROBELKYLK 540
 Qy 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFAVSLAFNILREPLNLP 600
 Db 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFAVSLAFNILREPLNLP 600
 Qy 601 MVLSIIYOASVSLRLRIPLSHELEPDSIERRPVKGCGTNSITTVNATTTARSPPT 660
 Db 601 MVLSIIYOASVSLRLRIPLSHELEPDSIERRPVKGCGTNSITTVNATTTARSPPT 660
 Qy 661 LMGITFESIPEGALVAVVQVCGCKSSLSALLAEMDKVGHVAKGSVAVYPOQAMIOND 720
 Db 661 LMGITFESIPEGALVAVVQVCGCKSSLSALLAEMDKVGHVAKGSVAVYPOQAMIOND 720
 Qy 721 SLRENILFGCQLEPPYRSVIOACALLPDLILPSGRTIEGKGVNLGGOKQVSLAR 780
 Db 721 SLRENILFGCQLEPPYRSVIOACALLPDLILPSGRTIEGKGVNLGGOKQVSLAR 780
 Qy 781 AVVSNADIYFDDPLSVDVAHVGHIFENYIGPVGMLKNTKRIIVTHSMSTLPDQVYIV 840
 Db 781 AVVSNADIYFDDPLSVDVAHVGHIFENYIGPVGMLKNTKRIIVTHSMSTLPDQVYIV 840
 Qy 841 MSGGKISEMGSYOELARDGAFAEFLRTYASTEOEDAEENGVTGVSOPGEAKOMENGM 900
 Db 841 MSGGKISEMGSYOELARDGAFAEFLRTYASTEOEDAEENGVTGVSOPGEAKOMENGM 900
 Qy 901 LVYDSACKQLOROLSSSSSYSGDISRRHNSTAELOKAEKKEETWKLMEADKAQTQVKL 960
 Db 901 LVYDSACKQLOROLSSSSSYSGDISRRHNSTAELOKAEKKEETWKLMEADKAQTQVKL 960
 Qy 961 SYVWDYKKAIGLFSFLSIFLPMCNHVSALASNTWLSLMDDPVNGTOHTKRLSVYG 1020
 Db 961 SYVWDYKKAIGLFSFLSIFLPMCNHVSALASNTWLSLMDDPVNGTOHTKRLSVYG 1020
 Qy 1021 ALGISOGIAVFGYSMAVISGIIILASRCLHYDLHSILRSPMSFEERIPSGNLVVRFSKEL 1080
 Db 1021 ALGISOGIAVFGYSMAVISGIIILASRCLHYDLHSILRSPMSFEERIPSGNLVVRFSKEL 1080
 Qy 1081 DTYDSMIPEVIKMFMSGLFNVIKACIYILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
 Db 1081 DTYDSMIPEVIKMFMSGLFNVIKACIYILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
 Qy 1141 KRLESVRSRPSVYSHFNFTLLGVSTIRAFEBEORFIHOSDKVDENQKAYPSIYANMWA 1200
 Db 1141 KRLESVRSRPSVYSHFNFTLLGVSTIRAFEBEORFIHOSDKVDENQKAYPSIYANMWA 1200
 Qy 1201 VRLEVCVNCIYLFALFPAVISRHSLSAGLVGLSVSYLSQVTTYTNMIVRMSSEMETIIVA 1260
 Db 1201 VRLEVCVNCIYLFALFPAVISRHSLSAGLVGLSVSYLSQVTTYTNMIVRMSSEMETIIVA 1260
 Qy 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFRNYCLRYREDLDFVLRIHINVTINGG 1320
 Db 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFRNYCLRYREDLDFVLRIHINVTINGG 1320
 Qy 1321 EKVGIIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLAFKTIITIPQDVLFL 1380
 Db 1321 EKVGIIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLAFKTIITIPQDVLFL 1380
 Qy 1381 SCSLRNMLDPFSQYDEEVTSLLEALIKDFVSALPKLDHECEGEGENISVGOQROVLCL 1440
 Db 1381 SCSLRNMLDPFSQYDEEVTSLLEALIKDFVSALPKLDHECEGEGENISVGOQROVLCL 1440
 Qy 1441 ARALLRKTKLIVDEATAVAVDELTDDLIQSTIRTOFEDCTVLITAHRLANTIMDYTRIVL 1500
 Db 1441 ARALLRKTKLIVDEATAVAVDELTDDLIQSTIRTOFEDCTVLITAHRLANTIMDYTRIVL 1500

Qy 1501 DKGEIOEYGA PSDILOORGLFYSMAKDAGLVGGGGGMLSRKGIIPREYVLTRLAEPAE 1560
 Db 1501 DKGEIOEYGA PSDILOORGLFYSMAKDAGLVGGGGGMLSRKGIIPREYVLTRLAEPAE 1560
 Qy 1561 PRYTRERRRARFVSKKNCVNAHKNIREOGRFLQDVFPTTVLDLKMPTLLIFTMSFLCSW 1620
 Db 1561 PRYTRERRRARFVSKKNCVNAHKNIREOGRFLQDVFPTTVLDLKMPTLLIFTMSFLCSW 1620
 Qy 1621 LLEFAMVWLIAFAHG-----DLAP-----GEGNVPCVTSIHSFSSAFLF 1680
 Db 1621 LLEFAMVWLIAFAHG-----DLAP-----GEGNVPCVTSIHSFSSAFLF 1680
 Qy 1661 SIEVQVTIGRGRNVTEECPLATILILVQNVGLMNAIMLGCIFMTQAQHRAEFTLIF 1720
 Db 1661 SIEVQVTIGRGRNVTEECPLATILILVQNVGLMNAIMLGCIFMTQAQHRAEFTLIF 1720
 Qy 1721 SKIAVITLRHGRCLCFMLRVGDLRKSMIISATIHQVVRKTTSPGEVVPPLHQVDIPMENG 1780
 Db 1721 SKIAVITLRHGRCLCFMLRVGDLRKSMIISATIHQVVRKTTSPGEVVPPLHQVDIPMENG 1780
 Qy 1741 VGGNGIFLVAPLIIVHVIDSNSPLYDLAPSDLHHODLEIIVILEGVETTGITTOARTS 1800
 Db 1741 VGGNGIFLVAPLIIVHVIDSNSPLYDLAPSDLHHODLEIIVILEGVETTGITTOARTS 1800
 Qy 1841 YLADETLMGQRFVPIVAEEDGRYSVDYSKRGNTIKVPTPLCTARQLDEDSILDALTLAS 1900
 Db 1841 YLADETLMGQRFVPIVAEEDGRYSVDYSKRGNTIKVPTPLCTARQLDEDSILDALTLAS 1900
 Qy 1861 YLADETLMGQRFVPIVAEEDGRYSVDYSKRGNTIKVPTPLCTARQLDEDSILDALTLAS 1920
 Db 1861 YLADETLMGQRFVPIVAEEDGRYSVDYSKRGNTIKVPTPLCTARQLDEDSILDALTLAS 1920
 Qy 1901 SRGPLRKRSVAVAKAKPKFSISPDLS 1927
 Db 1901 SRGPLRKRSVAVAKAKPKFSISPDLS 1927
 Qy 1921 SRGPLRKRSVAVAKAKPKFSISPDLS 1947
 Db 1921 SRGPLRKRSVAVAKAKPKFSISPDLS 1947
 RESULT 4
 ADY86941
 ID ADY86941 standard; protein; 1891 AA.
 XX
 AC ADY86941;
 XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Human MRP1-Kir6.2 mutant delacta36 fusion protein, SEQ ID NO: 4.
 XX
 KM Ionophore; biosensor; drug screening; diagnostic;
 KM microorganism detection; potassium channel; fusion protein;
 KM multidrug resistance protein 1; MRP1; Kir6.2; murein.
 OS Homo sapiens.
 OS Mus musculus.
 OS Chimeric.
 XX
 FT Key
 FT Region location/Qualifiers
 FT /note="Multidrug resistance protein 1 (MRP1)"
 FT Region 1532..1537
 FT /note="Hexaglycine spacer"
 FT Region 1538..1891
 FT /note="Kir6.2 mutant delacta36 protein"
 PN US2005063989-A1.
 XX
 XX 24-MAR-2005.
 PD
 XX
 PF 22-SEP-2003; 2003US-00665283.
 PF
 PR 22-SEP-2003; 2003US-00665283.
 PR
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 XX
 XX Derand R, Garcia E, Prost A, Revilloud J, Viaudou M;
 PI
 XX WPI; 2005-252611/26.
 DR

XX New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
PS Claim 25; SEQ ID NO 4; 78pp; English.
CC The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 mutant detac36 protein.
XX
SQ Sequence 1891 AA;
Query Match 98.3%; Score 9734; DB 9; Length 1891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALGFCSDGSDPLMDNNTWNTNSNPDFTKCFQNTVWVPCFYLACFPFFLYLSRH 60
Db 1 MALGFCSDGSDPLMDNNTWNTNSNPDFTKCFQNTVWVPCFYLACFPFFLYLSRH 60
QY 61 DRGYIOMTPLNKTATAGFLIMIVCMADLFYSFWERSNGIFLAPVFLVSPILLGITTLA 120
Db 61 DRGYIOMTPLNKTATAGFLIMIVCMADLFYSFWERSNGIFLAPVFLVSPILLGITTLA 120
QY 121 TFLIQLEERRKGVSSGIMLTFWLVALVCALALRSKIMTALKEDAQVDLFDITFYYVS 180
Db 121 TFLIQLEERRKGVSSGIMLTFWLVALVCALALRSKIMTALKEDAQVDLFDITFYYVS 180
QY 181 LLLIQVLVSCSDSSRPFSETHDPNRPBESSASLSTITFWITGLVGRGROLESBD 240
Db 181 LLLIQVLVSCSDSSRPFSETHDPNRPBESSASLSTITFWITGLVGRGROLESBD 240
QY 241 LMSLNKEDTSQOVVPLVKNMKKECAKTRKOPVKVYSSKDPAPKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSQOVVPLVKNMKKECAKTRKOPVKVYSSKDPAPKESSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLYKTFGPYPLMSFFPKAIDLMFSGPOLIKLIFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFKVLYKTFGPYPLMSFFPKAIDLMFSGPOLIKLIFVNDTKAPD 360
QY 361 WQGYFYTVLLFVTACLOTLVLAHQYFHICFVSGMRIKTAVIGAVYRKALVITNSAKSSSTV 420
Db 361 WQGYFYTVLLFVTACLOTLVLAHQYFHICFVSGMRIKTAVIGAVYRKALVITNSAKSSSTV 420
QY 421 GEIYNLMSVDARFMDLATYINMISAPLOVIALYLMNLGSPVLAGVAVMVLMPVN 480
Db 421 GEIYNLMSVDARFMDLATYINMISAPLOVIALYLMNLGSPVLAGVAVMVLMPVN 480
QY 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKLVAMELAFQKVLAIROEBLKVLK 540
Db 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKLVAMELAFQKVLAIROEBLKVLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVTTDENNILLDAQTAFFVSLAFNIIIRPLANTLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVTTDENNILLDAQTAFFVSLAFNIIIRPLANTLP 600
QY 601 MVISIVQASVSLKRLIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDDEPT 660
Db 601 MVISIVQASVSLKRLIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDDEPT 660
QY 661 LAGTTFSPREGALVAVGVOGCKSSLSALLAEMDKYEGHVAIKGSVAVYVQQAIOND 720
Db 661 LAGTTFSPREGALVAVGVOGCKSSLSALLAEMDKYEGHVAIKGSVAVYVQQAIOND 720
QY 721 SLRENILFGQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGGOKORVSLAR 780
Db 721 SLRENILFGQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGGOKORVSLAR 780

Db 721 SLRENILFGQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGGOKORVSLAR 780
QY 781 AVYSNADIYLEDPLSAVDAAVHGKIFENVIGPKMLKNKTRIIIVTHSMSTYLPQVDVIV 840
Db 781 AVYSNADIYLEDPLSAVDAAVHGKIFENVIGPKMLKNKTRIIIVTHSMSTYLPQVDVIV 840
QY 841 MSGGKISEMGSYQGLLARDGAFAPFLTYASTBEQDAAEENGVTGSGPGCEAAQOMENG 900
Db 841 MSGGKISEMGSYQGLLARDGAFAPFLTYASTBEQDAAEENGVTGSGPGCEAAQOMENG 900
QY 901 LVTSAGKQLOROUSSSSSYSGDISRHNNSTAELOKAEKKEETWKLMEADKAQOQVYL 960
Db 901 LVTSAGKQLOROUSSSSSYSGDISRHNNSTAELOKAEKKEETWKLMEADKAQOQVYL 960
QY 961 SVYWDYKAIQGLFISPIFLFMCNVAASALSNYMLSLMTDDPIVNGTQEHKTVRLSYG 1020
Db 961 SVYWDYKAIQGLFISPIFLFMCNVAASALSNYMLSLMTDDPIVNGTQEHKTVRLSYG 1020
QY 1021 ALGISQGIYAVRGYMANYSIGGILASRCLHVDLHSLRSPMSFFERTPSGNLVNRFSEL 1080
Db 1021 ALGISQGIYAVRGYMANYSIGGILASRCLHVDLHSLRSPMSFFERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMIPEVTKMFMGSLFENVIGACIVILLATPIAAIIIPPLGLIYFVQGFYVASSRQL 1140
Db 1081 DTVDSMIPEVTKMFMGSLFENVIGACIVILLATPIAAIIIPPLGLIYFVQGFYVASSRQL 1140
QY 1141 KRLSEVSRSPPYSHFNETLLGVSVIRAFEEQERFIHOSDLKYDENOKAYPSIVANRWLA 1200
Db 1141 KRLSEVSRSPPYSHFNETLLGVSVIRAFEEQERFIHOSDLKYDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCTYLPALPAVIRSHLSAGLVLSYSIQVTTYINMLVRMSSEMETNIVA 1260
Db 1201 VRLCEVGNCTYLPALPAVIRSHLSAGLVLSYSIQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEVSETEKEAPMOIQETAPPSMPOQVGRVFRNYCLRYREDLDFVLSHINVTINGG 1320
Db 1261 VERLKEVSETEKEAPMOIQETAPPSMPOQVGRVFRNYCLRYREDLDFVLSHINVTINGG 1320
QY 1321 EKVGIYGRGTGAKSSLTLGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIIPDPVLF 1380
Db 1321 EKVGIYGRGTGAKSSLTLGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIIPDPVLF 1380
QY 1381 SGSLRMLNDPFSQYSDSEEWNTSLAHKDVVSALPDLDEHCEAGGNNLSVGORQLVCL 1440
Db 1381 SGSLRMLNDPFSQYSDSEEWNTSLAHKDVVSALPDLDEHCEAGGNNLSVGORQLVCL 1440
QY 1441 ABALRKTKIIVLEAATAVDETDLLQSTIFRFQFEDCTVLTAAHRLNTIMDYTRVYL 1500
Db 1441 ABALRKTKIIVLEAATAVDETDLLQSTIFRFQFEDCTVLTAAHRLNTIMDYTRVYL 1500
QY 1501 DKGEIOEYGAESDILLQGRGLFYMAKQAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
Db 1501 DKGEIOEYGAESDILLQGRGLFYMAKQAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
QY 1561 PRYTRERRARFVSKGNCNVAHKNIREQGFLOVFTTLYDLKXPHULLFTMSFLCSW 1620
Db 1561 PRYTRERRARFVSKGNCNVAHKNIREQGFLOVFTTLYDLKXPHULLFTMSFLCSW 1620
QY 1621 LLFAMVWMLIAPAGDGLAPEGSTNVCVTSIHSSSAFLFSIEVOVTTIGFGGRMTEBCP 1680
Db 1621 LLFAMVWMLIAPAGDGLAPEGSTNVCVTSIHSSSAFLFSIEVOVTTIGFGGRMTEBCP 1680
QY 1681 LAIILIVQNTVGLMINAIMGCIFFMKTQAHRRAETLIFSKHAVITTLRHGLCEMLRVG 1740
Db 1681 LAIILIVQNTVGLMINAIMGCIFFMKTQAHRRAETLIFSKHAVITTLRHGLCEMLRVG 1740
QY 1741 DLKSMIISATIHQVAKTSSPEGEVPLHQVDPIMENGVGNGIFLVAPLIIYHVIDS 1800
Db 1741 DLKSMIISATIHQVAKTSSPEGEVPLHQVDPIMENGVGNGIFLVAPLIIYHVIDS 1800
QY 1801 NSPLYDLAPSDLHHQDLEITVILEGVETTGITTOATSVYLADEIILMGGORFVPIVAED 1860
Db 1801 NSPLYDLAPSDLHHQDLEITVILEGVETTGITTOATSVYLADEIILMGGORFVPIVAED 1860

[illegible]

Db	61	DRGIIQMTPLNKTKTALGFLMIVICADLHYSEWESRGLFLPVLVSFTLGLITTLA	120
Qy	121	TFLLQERBKVOSSGIMLTFMVALVACALAIRSKIMTALKEBDAQVDLFRTITFYVYS	180
Db	121	TFLLQERBKVOSSGIMLTFMVALVACALAIRSKIMTALKEBDAQVDLFRTITFYVYS	180
Qy	181	LLLIQVLVSCSDRSLPFSETHIDPNCPBSSASPLSRITFMWITGLIVRGYQPLEGSD	240
Db	181	LLLIQVLVSCSDRSLPFSETHIDPNCPBSSASPLSRITFMWITGLIVRGYQPLEGSD	240
Qy	241	LMSLNKEDTSQOVVPUTVKRWKKECATRQOPKVVVYSSDPAQPKSSSYVDANEVEAL	300
Db	241	LMSLNKEDTSQOVVPUTVKRWKKECATRQOPKVVVYSSDPAQPKSSSYVDANEVEAL	300
Qy	301	IVKSPQEMKPSLEKVLVYKTFGPFLMSFFFKAIHIDLMFSGQIILKLIKFNDRKAPD	360
Db	301	IVKSPQEMKPSLEKVLVYKTFGPFLMSFFFKAIHIDLMFSGQIILKLIKFNDRKAPD	360
Qy	361	WQGYFTYVLLFVTAQLOTLVLHOFPHICFVSGNRKITAIVIGAYRKALVITNSARSGSTV	420
Db	361	WQGYFTYVLLFVTAQLOTLVLHOFPHICFVSGNRKITAIVIGAYRKALVITNSARSGSTV	420
Qy	421	GEIYNLMSVDAQRPMDLATTYNNIWSAPLOVITLILMLNIGPSYLAGVAVMTLVPNV	480
Db	421	GEIYNLMSVDAQRPMDLATTYNNIWSAPLOVITLILMLNIGPSYLAGVAVMTLVPNV	480
Qy	481	AVMMKRTKYOVAMHMSKONRIKLMELINGIVLKLVMELAFKOKVLAIRQELKVLK	540
Db	481	AVMMKRTKYOVAMHMSKONRIKLMELINGIVLKLVMELAFKOKVLAIRQELKVLK	540
Qy	541	KSAYLSAVGTFTWCTPEFLVACTPFAVYVITIDENNILDAQTAAPVSLATFNIRLFPNTILP	600
Db	541	KSAYLSAVGTFTWCTPEFLVACTPFAVYVITIDENNILDAQTAAPVSLATFNIRLFPNTILP	600
Qy	601	MWISSIVOASVSLKRLRILFLSHELEPDSIERRPVKDGGGTNSITVRNATFTWARSDEPT	660
Db	601	MWISSIVOASVSLKRLRILFLSHELEPDSIERRPVKDGGGTNSITVRNATFTWARSDEPT	660
Qy	661	LNCTFTSIPREGALVAVVGVGVGCKSSLSLALAEMLKVGHAIAKGSVAVVPOAWIQND	720
Db	661	LNCTFTSIPREGALVAVVGVGVGCKSSLSLALAEMLKVGHAIAKGSVAVVPOAWIQND	720
Qy	721	SLRENIIFGQOLEBPYRSYIOACALLPDEILPQSDRTEIGKGNLSGGOKORSYLSLR	780
Db	721	SLRENIIFGQOLEBPYRSYIOACALLPDEILPQSDRTEIGKGNLSGGOKORSYLSLR	780
Qy	781	AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLKXKTRILVTHSMSYLPOVDVILV	840
Db	781	AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLKXKTRILVTHSMSYLPOVDVILV	840
Qy	841	MSGGKISEMGSYOBELARDGAFMEFLRTTASTQEOBDAENGVYGSFGKEAKOMENGM	900
Db	841	MSGGKISEMGSYOBELARDGAFMEFLRTTASTQEOBDAENGVYGSFGKEAKOMENGM	900
Qy	901	LYTDSAGKOLQRLTSSSSSYSGDISRHNSHTLQLOAKKKEBTMLKMDADRAQGVQL	960
Db	901	LYTDSAGKOLQRLTSSSSSYSGDISRHNSHTLQLOAKKKEBTMLKMDADRAQGVQL	960
Qy	961	SVYWDYKACIGLEISFLSIFLPMCHNVSLASNYWMLSTWDDPIVNGTOEHKRYLSYVG	1020
Db	961	SVYWDYKACIGLEISFLSIFLPMCHNVSLASNYWMLSTWDDPIVNGTOEHKRYLSYVG	1020
Qy	1021	ALGISQIAVFGYSMAVSTOGLIASRCLHVDLHSLTNSPMGFPEETPSGNLVNPSKEL	1080
Db	1021	ALGISQIAVFGYSMAVSTOGLIASRCLHVDLHSLTNSPMGFPEETPSGNLVNPSKEL	1080
Qy	1081	DTYDSMTPEYIKMFMSSLFNVIGACVILLATPIAIIIPGLIYFPPQRYVASSRQL	1140
Db	1081	DTYDSMTPEYIKMFMSSLFNVIGACVILLATPIAIIIPGLIYFPPQRYVASSRQL	1140
Qy	1141	KRLSEVSRSPPVSHFNETLLGVSVIRAFEEQERFIHQSDLKVQENOKAYYPSIVANRWLA	1200

Db 1141 KRLSEVSRSPVYSHPNETLGLSVIRAEEDERFIHQSDLKVDENOKAYRYSIVANKRLA 1200
 QY 1201 VRLCEVCANCIVLFAALFAVISRHSISAGLVGISVSYIQVTTYINMLVRMSSEMETINVA 1260
 Db 1201 VRLCEVCANCIVLFAALFAVISRHSISAGLVGISVSYIQVTTYINMLVRMSSEMETINVA 1260
 QY 1261 VERLKEYSETEKPAWQIQETAPRPSMWQVGRVFRANCLAREBDLVLRHINVTINGG 1320
 Db 1261 VERLKEYSETEKPAWQIQETAPRPSMWQVGRVFRANCLAREBDLVLRHINVTINGG 1320
 QY 1321 EKVGIVRTGAGKSLTGLFRINSEAGEIIIDININAKIGLHDLRFKTIIPDDPYLF 1380
 Db 1321 EKVGIVRTGAGKSLTGLFRINSEAGEIIIDININAKIGLHDLRFKTIIPDDPYLF 1380
 QY 1381 SGLSRMNLDPFSSQYSDSEEWTSLELAHLKDFVSALPDLHDECAAGGENTLSVGORQLVCL 1440
 Db 1381 SGLSRMNLDPFSSQYSDSEEWTSLELAHLKDFVSALPDLHDECAAGGENTLSVGORQLVCL 1440
 QY 1441 ARALLRKTILVDEBATAVVDLETDLIQSTIRTFQFECTVLTIAHRLNTIMDYTRIVVL 1500
 Db 1441 ARALLRKTILVDEBATAVVDLETDLIQSTIRTFQFECTVLTIAHRLNTIMDYTRIVVL 1500
 QY 1501 DKGEIOEXGAPSDLLQORGLFYMAKDAGLV 1531
 Db 1501 DKGEIOEXGAPSDLLQORGLFYMAKDAGLV 1531

RESULT 6
 AAM74471
 ID AAM74471 standard; protein; 1531 AA.

AC AAM74471;
 DT 18-MAY-1999 (first entry)

XX Human multidrug resistance-associated protein variant.

XX Multidrug resistance-associated protein; MDR; human; diagnosis;
 KM MDR tumour cell identification; cancer therapy.

OS Homo sapiens.

XX US5882875-A.

PD 16-MAR-1999.

XX 05-JUN-1995; 95US-00462109.

XX 27-OCT-1992; 92US-00966923.

XX 08-MAR-1993; 93US-00029340.

XX 26-OCT-1993; 93US-00141893.

XX 20-MAR-1995; 95US-00407207.

XX (TMOH) UNIV QUEBENS KINGSTON.

XX Cole SPC, Deeley RG;
 WI: 1999-214061/18.

XX N-PSDB; AAX21977.

PT Identifying a multidrug resistant tumor cell by contacting the cell with
 an antibody/antigen-binding fragment - which binds to an expressed
 protein encoded by multidrug resistance-associated protein (MRP) nucleic
 acid.

PS Claim 3; Col 69-80; 80pp; English.

CC This sequence is the human multidrug resistance-associated (MDR) protein.
 CC The invention relates to a method for identifying a multidrug resistant
 CC (MDR) tumor cell. Compositions and methods utilizing the MDR proteins
 CC can be used to treat patients with tumours displaying multidrug
 CC resistance, particularly those displaying resistance to antineoplastic
 CC epipodophyllotoxins, vinca alkaloids, and hydrophobic drugs. The methods

CC for inhibiting/killing a MDR tumor cell can be useful for treating
 CC breast cancer, leukemias, fibrosarcomas, cervical cancer, gliomas,
 CC thymomas, neuroblastomas and lung cancer. The MDR DNA sequence when
 CC labeled are useful as molecular probes for diagnosing multidrug
 CC resistance of a tumor (using cells from a tumour biopsy) and for
 CC designing ribozymes which are capable of cleaving a single-stranded
 CC nucleic acid encoding a protein having MRP activity. Recombinant
 CC expression vectors containing human MDR coding sequences can be
 CC transfected into a drug sensitive cell line to produce a protein in the
 CC cell which confers MDR, protecting non-resistant non-tumour cells from
 CC the effects of chemotherapeutics has major clinical importance. Cells
 CC transformed with the MDR coding sequences are useful for testing
 CC potential therapeutic agents for their effectiveness against MDR cells
 CC and for identifying chemosensitizers of a therapeutic agent

CC Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 2; Length 1531;
 Best local similarity 100.0%; Pred. NO. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNVMTNSNDFPKFQNTVLWVPCFYLMACFPPELYLSRH 60
 Db 1 MALRGFCSADGSDPLMDMNVMTNSNDFPKFQNTVLWVPCFYLMACFPPELYLSRH 60
 QY 61 DRGYIQMTPLNKTKTALGFLIMVCMADLFYSFWEKSRGIFLAVFLVSPTLGITTLA 120
 Db 61 DRGYIQMTPLNKTKTALGFLIMVCMADLFYSFWEKSRGIFLAVFLVSPTLGITTLA 120
 QY 121 TELLOERRKGVSSGIMLTFMLVALCALALRSKIMTAKEDAOVDLFPDIFPYVFS 180
 Db 121 TELLOERRKGVSSGIMLTFMLVALCALALRSKIMTAKEDAOVDLFPDIFPYVFS 180
 QY 181 LLLIQVLVSCFSDSPFLFSETIHDNPNCPSESSAFLSRITFWITGLIVRGYROPLEBSD 240
 Db 181 LLLIQVLVSCFSDSPFLFSETIHDNPNCPSESSAFLSRITFWITGLIVRGYROPLEBSD 240
 QY 241 LMSINKEDTSFOVVPVIVKMKKECAKTRKOPVAVVYSSKDPAPKRESSKYDANEVEAL 300
 Db 241 LMSINKEDTSFOVVPVIVKMKKECAKTRKOPVAVVYSSKDPAPKRESSKYDANEVEAL 300
 QY 301 IVKSPKEMNPSLKKULYKTFGPFLMSFPFKALHDLMMFSGPOLIKLILFVNDTKAPD 360
 Db 301 IVKSPKEMNPSLKKULYKTFGPFLMSFPFKALHDLMMFSGPOLIKLILFVNDTKAPD 360
 QY 361 WQGYFYTVLFPVTACLOTLYVHOYFHCFSVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
 Db 361 WQGYFYTVLFPVTACLOTLYVHOYFHCFSVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
 QY 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOVITALLYLMLNGLSPVLAGVAVMYLMEVFN 480
 Db 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOVITALLYLMLNGLSPVLAGVAVMYLMEVFN 480
 QY 481 AVMAKTKTYOVAMKSKDNRIKLMNELINIKYLUKAYAMELAKDKYLAEROELKYLK 540
 Db 481 AVMAKTKTYOVAMKSKDNRIKLMNELINIKYLUKAYAMELAKDKYLAEROELKYLK 540
 QY 541 KSAVLSAVGTFPTWCTPPLVALCTPAYVVTIDENNILDAQTAFFSALFNILRPPLNTLP 600
 Db 541 KSAVLSAVGTFPTWCTPPLVALCTPAYVVTIDENNILDAQTAFFSALFNILRPPLNTLP 600
 QY 601 MVISIVQASVSLKRLRIFLSHELEPDSIRRPVKOGGINSITVRNATFTWASDPT 660
 Db 601 MVISIVQASVSLKRLRIFLSHELEPDSIRRPVKOGGINSITVRNATFTWASDPT 660
 QY 661 LINGTFTSIPBEGALVAVVQVCGKSSLLSALLAEMDKYEGVAIKGSVAVYPOQAMTQND 720
 Db 661 LINGTFTSIPBEGALVAVVQVCGKSSLLSALLAEMDKYEGVAIKGSVAVYPOQAMTQND 720
 QY 721 SLRENIILFGCOLPEPYRSVVOACALPDLIELPSGDRTEIGEGVNLSSGGOKORVSLAR 780
 Db 721 SLRENIILFGCOLPEPYRSVVOACALPDLIELPSGDRTEIGEGVNLSSGGOKORVSLAR 780

QY 781 AVYSNADITYLFDPLSAVDAAHGKHI FENVIGPKMKLNKTRILVTHSMYLPQVDYIIV 840
Db 781 AVYSNADITYLFDPLSAVDAAHGKHI FENVIGPKMKLNKTRILVTHSMYLPQVDYIIV 840
QY 841 MSGGKISMSGSYQELLADGAFAEFLRTYASTEQDABENGVTYVSGPGKAKOMENGM 900
Db 841 MSGGKISMSGSYQELLADGAFAEFLRTYASTEQDABENGVTYVSGPGKAKOMENGM 900
QY 901 LVTSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKAKKEETMKLEADAKAOTGQYKL 960
Db 901 LVTSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKAKKEETMKLEADAKAOTGQYKL 960
QY 961 SVYDYMKAIGLFIISFLSIPLFMCHNVSAALSNVYLSMTDDPIVNGTQEHTKYRLSYG 1020
Db 961 SVYDYMKAIGLFIISFLSIPLFMCHNVSAALSNVYLSMTDDPIVNGTQEHTKYRLSYG 1020
QY 1021 ALGISQGIYAVFGYMAVSIIGGILASRCLHVDLHLSILSPMSFFERTPSGNLVNRFSEL 1080
Db 1021 ALGISQGIYAVFGYMAVSIIGGILASRCLHVDLHLSILSPMSFFERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMIPEVIKMFMSGLFENVIGACIVILLATPIAIIIPPLGIIYFFVORFVYASSROL 1140
Db 1081 DTVDSMIPEVIKMFMSGLFENVIGACIVILLATPIAIIIPPLGIIYFFVORFVYASSROL 1140
QY 1141 KRLESVSRSPYVSHFNETLGVSVLRAFEQERFIHQDLKVDENOKAYPSIYANRWLA 1200
Db 1141 KRLESVSRSPYVSHFNETLGVSVLRAFEQERFIHQDLKVDENOKAYPSIYANRWLA 1200
QY 1201 VRLCEVNGCIVLPALFAVIRSHSISAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
Db 1201 VRLCEVNGCIVLPALFAVIRSHSISAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLEKESVETKEKAPWQIQETAPPSWPQNGRVERRANCLAFREYDLVDVLAHINVTINGG 1320
Db 1261 VERLEKESVETKEKAPWQIQETAPPSWPQNGRVERRANCLAFREYDLVDVLAHINVTINGG 1320
QY 1321 EKVGIIVGTGAGKASLTGLFRINESAGEFIIIDGINAKI GLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIIVGTGAGKASLTGLFRINESAGEFIIIDGINAKI GLHDLRFKTIIPQDPVLF 1380
QY 1381 SGIJRMNLDPPSQYSDDEEWTSLBLAHLKDFVSALPDKLDEHCAEGENLSVGORQLVCL 1440
Db 1381 SGIJRMNLDPPSQYSDDEEWTSLBLAHLKDFVSALPDKLDEHCAEGENLSVGORQLVCL 1440
QY 1441 ARALLRKTKIIVDEATAVULETDDLIQSTIRPOFECYTLTAHRLNTIMDTTRYIVL 1500
Db 1441 ARALLRKTKIIVDEATAVULETDDLIQSTIRPOFECYTLTAHRLNTIMDTTRYIVL 1500
QY 1501 DKGEIOEGAPSDLLQQRGLFYMAKADAGLV 1531
Db 1501 DKGEIOEGAPSDLLQQRGLFYMAKADAGLV 1531

RESULT 7
AAW9894
ID AAW9894 standard; protein; 1531 AA.
XX AAW9894;
AC
DT 10-JUN-1999 (first entry)
XX
DE Human multidrug resistance-associated protein natural variant.
XX
KM Human; multidrug resistance-associated protein; MRP; cytotoxic drug;
XX cancer; chemotherapy.
XX
OS Homo sapiens.
XX
PN US5891724-A.
XX
PD 06-APR-1999.
XX
PF 05-JUN-1995; 95US-00460907.

XX 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
PA (TOOH) UNIV QUEBENS KINGSTON.
XX
PI Cole SPC, Dealey RG;
XX
XX WPI; 1999-253868/21.
DR N-PSDB; AAX19818.
DR
XX
PT Protecting mammalian cells against cytotoxic drugs.
PS
PS Claim 2; Col 79-86; 82pp; English.
XX
XX The present sequence represents a human multidrug resistance-associated
CC protein (MRP). The present invention also describes a method for
CC protecting a mammalian cell against the cytotoxicity of anthracyclines,
CC epipodophyllotoxins and vinca alkaloids (A) by introducing into it a
CC nucleic acid (I) that hybridizes under stringent conditions to a nucleic
CC acid (II) that encodes an MRP protein (III). Introduction of (I) protects
CC cells against cytotoxic effects of (A), particularly to protect normal
CC cells against (A) being used for treatment of cancers. Cells transformed
CC with (I) can be used to screen for agents that affect multidrug
CC resistance or are directly toxic to multidrug resistant cells, i.e.
CC potential therapeutics for multidrug-resistant cancers. Confering
CC resistance to normal cells should allow an increase in the dose of (A)
CC that can be administered safely
XX
SQ Sequence 1531 AA:
Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLDMNMTWNTSNPDKCFQNTLVWVPCPYLACPFYFLYLSRH 60
Db 1 MALRGFCSADGSDPLDMNMTWNTSNPDKCFQNTLVWVPCPYLACPFYFLYLSRH 60
QY 61 DRGYQMTPLNKTALGFLMIVCWADLFYSEFERSRGIFLAPVFLVSPILIGITLLA 120
Db 61 DRGYQMTPLNKTALGFLMIVCWADLFYSEFERSRGIFLAPVFLVSPILIGITLLA 120
QY 121 TFLIQERRRGVSSGIMLTFWLVAVCALAIRSKIMTALKEDAOYDLFRDITFYVYS 180
Db 121 TFLIQERRRGVSSGIMLTFWLVAVCALAIRSKIMTALKEDAOYDLFRDITFYVYS 180
QY 181 LLLIQVLSCGSPRSPLFSTIHDNPNCPRESSASFLSRITFWMTGLIVRGYRPLEGSD 240
Db 181 LLLIQVLSCGSPRSPLFSTIHDNPNCPRESSASFLSRITFWMTGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVVKMKKECAKTRKOPVYVSSKOPAPKESKRYDANEVEAL 300
Db 241 LMSLNKEDTSEOVVPLVVKMKKECAKTRKOPVYVSSKOPAPKESKRYDANEVEAL 300
QY 301 IVKSPQKEMNPSEKVLTKTFGPFLMSFFPKAIHDLMEGPDILKLIKFVNDTRAPD 360
Db 301 IVKSPQKEMNPSEKVLTKTFGPFLMSFFPKAIHDLMEGPDILKLIKFVNDTRAPD 360
QY 361 WQGYFYVLLFVTRACLOTVLVLAQYFHCIFVSGMRKIKTAVYGAAYRKALVTTNSARKSSTV 420
Db 361 WQGYFYVLLFVTRACLOTVLVLAQYFHCIFVSGMRKIKTAVYGAAYRKALVTTNSARKSSTV 420
QY 421 GEIVNLSVDAORFMDLATYINMIWSAPLGVIATLYLMLNLPSPVLAGVAVWVLPVNV 480
Db 421 GEIVNLSVDAORFMDLATYINMIWSAPLGVIATLYLMLNLPSPVLAGVAVWVLPVNV 480
QY 481 AVNAAKTKTYQVAHMSKDNRIKLMNEILNGIRVLYKLYAMELAKPKDKVLAIROBELKVLK 540
Db 481 AVNAAKTKTYQVAHMSKDNRIKLMNEILNGIRVLYKLYAMELAKPKDKVLAIROBELKVLK 540

```

QY 541 KSAVLSANGTWTWCTPPLVALCTPAVYVTTIDENNILDAQTAFFVSLAFLNIRFPLNLP 600
Db 541 KSAVLSANGTWTWCTPPLVALCTPAVYVTTIDENNILDAQTAFFVSLAFLNIRFPLNLP 600
QY 601 MVISSIVASVSLKRLIFLSHEELEPDSIERRPVKDGGNTSIVRNATFTWASDDPT 660
Db 601 MVISSIVASVSLKRLIFLSHEELEPDSIERRPVKDGGNTSIVRNATFTWASDDPT 660
QY 661 LINGITFSIPEGALVAVGVGVGCKSSLSALLAENDKYEGHVAIKGSVAVYVPOQAMIOND 720
Db 661 LINGITFSIPEGALVAVGVGVGCKSSLSALLAENDKYEGHVAIKGSVAVYVPOQAMIOND 720
QY 721 SLRENIIFGCQLEPEYRSVIOACALLPDLLETLPSGDRTEIGEKGVNISGGQKQVSLAR 780
Db 721 SLRENIIFGCQLEPEYRSVIOACALLPDLLETLPSGDRTEIGEKGVNISGGQKQVSLAR 780
QY 781 AVYSNADLYLPDDPLSAVDAAHVKHIFENVIGPKMKLNKRIIVTSHMSYLPQVDVILV 840
Db 781 AVYSNADLYLPDDPLSAVDAAHVKHIFENVIGPKMKLNKRIIVTSHMSYLPQVDVILV 840
QY 841 MSGKISEMGVYOELLARDGAFAEFLRTYASTEOBDAENGVTGSGPKKAKOMENG 900
Db 841 MSGKISEMGVYOELLARDGAFAEFLRTYASTEOBDAENGVTGSGPKKAKOMENG 900
QY 901 LVTTDSAGQOLOROLSSSSSSGDISRHNSTAELQKAEKKETWKLMEADKAOIGQVYL 960
Db 901 LVTTDSAGQOLOROLSSSSSSGDISRHNSTAELQKAEKKETWKLMEADKAOIGQVYL 960
QY 961 SVYMDYMAKIGLFIISFLIPMCMHVASLANSWLSLMTDPIVNGOEHKVALSYG 1020
Db 961 SVYMDYMAKIGLFIISFLIPMCMHVASLANSWLSLMTDPIVNGOEHKVALSYG 1020
QY 1021 ALGISQGIAPVGYMAVSIIGLILASRCLHVDLHLSILSPMSFFERTSGNLVNRFSKEL 1080
Db 1021 ALGISQGIAPVGYMAVSIIGLILASRCLHVDLHLSILSPMSFFERTSGNLVNRFSKEL 1080
QY 1081 DTVDSMIEBVIKMFNGSLFNVTGACIVILLATPIAIIIPPLGLYFFVQRPYVASSROL 1140
Db 1081 DTVDSMIEBVIKMFNGSLFNVTGACIVILLATPIAIIIPPLGLYFFVQRPYVASSROL 1140
QY 1141 KRLESVSRSPPVSHNETLLGVSVIRAFEBQERFHQSOLKVDENOKAYPBIIVANRWLA 1200
Db 1141 KRLESVSRSPPVSHNETLLGVSVIRAFEBQERFHQSOLKVDENOKAYPBIIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGSVSLQVTTYLNMLVMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGSVSLQVTTYLNMLVMSSEMETNIVA 1260
QY 1261 VERLKEVSETEKEAPWQIOETAPPSGWPQVGEFRNVCRLRYREDLDVYLRIHINTINGG 1320
Db 1261 VERLKEVSETEKEAPWQIOETAPPSGWPQVGEFRNVCRLRYREDLDVYLRIHINTINGG 1320
QY 1321 EKVGVIGRTGAGKSLTGLPRINSABEIIIDGINIAKIGLHDLRFITIIIPODPVLF 1380
Db 1321 EKVGVIGRTGAGKSLTGLPRINSABEIIIDGINIAKIGLHDLRFITIIIPODPVLF 1380
QY 1381 SGLSLMNLDPSFOYDEDEWMTSLFLAHLKDFVSALPDKLDBHCAEGENLSYGQOLVCL 1440
Db 1381 SGLSLMNLDPSFOYDEDEWMTSLFLAHLKDFVSALPDKLDBHCAEGENLSYGQOLVCL 1440
QY 1441 ARALIRKRIIVLDEATAVADLETDDLQSTIRTOFEDCTVLTIAHRLNTIMDYRVIVL 1500
Db 1441 ARALIRKRIIVLDEATAVADLETDDLQSTIRTOFEDCTVLTIAHRLNTIMDYRVIVL 1500
QY 1501 DKGEIQEYGAPSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIQEYGAPSDLLQORGLFYSMAKDAGLV 1531

```

```

AC AA55799;
XX 28-FEB-2000 (first entry)
XX Human multidrug resistance-associated protein (MRP) variant.
DE Chemosensitizer; multidrug resistance-associated protein; MRP; human;
XX therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;
XX cancer; variant.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc-difference 685
FT /label= L685S
FT /note= "wild-type Leu is replaced by Ser"
FT Misc-difference 1282
FT /label= R1282A
FT /note= "wild-type Arg is replaced by Ala"
XX
XX US6001563-A.
XX
XX 14-DEC-1999.
XX
XX 05-JUN-1995; 95US-00463179.
XX
XX 27-OCT-1992; 92US-00966923.
XX 08-MAR-1993; 93US-00029340.
XX 26-OCT-1993; 93US-00141893.
XX 20-MAR-1995; 95US-00407207.
XX
XX (TOOH ) UNIV QUEBENS KINGSTON.
XX
XX Cole SP, Deeley RG;
PI WPI: 2000-061877/05.
DR N-PDB; AA239556.
DR
PT Identification of chemosensitizers useful for treating cancer, using
XX nucleic acids encoding multidrug resistance-associated protein.
XX
XX Claim 3; Col 69-80; 77pp; English.
XX
XX The invention provides a method for identifying a substance which is a
XX chemosensitizer that comprises, contacting a cell transfected with
XX nucleic acid encoding multidrug resistance-associated protein (MRP) with
XX a therapeutic agent in vitro. The method is useful for identifying
XX chemosensitizers which may then be used to treat cancer (especially lung
XX cancer). The method allows the identification of chemosensitizers which
XX do not reverse P-glycoprotein-mediated multidrug resistance. The present
XX sequence represents a human MRP variant
XX
SQ Sequence 1531 AA:
Query Match 79.4%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRFGCADSDSDPLMDNNTWNTNSNPDFTCCFQNTVYVWPCFYLMACFPEFLYLSRH 60
Db 1 MALRFGCADSDSDPLMDNNTWNTNSNPDFTCCFQNTVYVWPCFYLMACFPEFLYLSRH 60
QY 61 DRGYQMPPLNKTKTALGFLIMVCMADLFYSFMRSGRIFLAPVFLVSPILLGITTLIA 120
Db 61 DRGYQMPPLNKTKTALGFLIMVCMADLFYSFMRSGRIFLAPVFLVSPILLGITTLIA 120
QY 121 TFLQLERKGVQSSGIMLTFMLVALVCALAILRSKIMTALKEDQVDLPFDITFYVYFS 180
Db 121 TFLQLERKGVQSSGIMLTFMLVALVCALAILRSKIMTALKEDQVDLPFDITFYVYFS 180
QY 181 LLLIQVLVSCFSDSDSPLFSETIHDNPPCPSSASGLSTITFTWITGLIVRGROPLRGSD 240
Db 181 LLLIQVLVSCFSDSDSPLFSETIHDNPPCPSSASGLSTITFTWITGLIVRGROPLRGSD 240

```

RESULT 8
 AA55799
 ID AA55799 standard; protein; 1531 AA.
 XX

QY 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPQESSKYDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPQESSKYDANEVEAL 300
 QY 301 IVKSPQKEMNPDLFKVLYKTFGPYFLMSFFPKAIDHLMFSGPOLIKLLIKFVNDTKAPD 360
 DB 301 IVKSPQKEMNPDLFKVLYKTFGPYFLMSFFPKAIDHLMFSGPOLIKLLIKFVNDTKAPD 360
 QY 361 WQGFYFVTLFVTLCTQTLVHOVPHICFVSGMRTKTAIVIGAVRKALVITNSARKSTV 420
 DB 361 WQGFYFVTLFVTLCTQTLVHOVPHICFVSGMRTKTAIVIGAVRKALVITNSARKSTV 420
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLMNLGSPVLGAVVWMLVNVN 480
 DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLMNLGSPVLGAVVWMLVNVN 480
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFDKVLAIROBELKVLK 540
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFDKVLAIROBELKVLK 540
 QY 541 KSAVLISAVGTFTWCTPPLVALCTPAVYVITIDENNILDAQTAFLVSLALFNILRPLNLP 600
 DB 541 KSAVLISAVGTFTWCTPPLVALCTPAVYVITIDENNILDAQTAFLVSLALFNILRPLNLP 600
 QY 601 MVISIIVQASVSLKRLRIFLSHEELEPDSIERRPYKDGSGNISTVRNATFTMARSDPT 660
 DB 601 MVISIIVQASVSLKRLRIFLSHEELEPDSIERRPYKDGSGNISTVRNATFTMARSDPT 660
 QY 661 LMGITFSIPGALVAVVQVCGSKSLLSALLAEMDKVGHVAKGSVAVYPQOAMTQND 720
 DB 661 LMGITFSIPGALVAVVQVCGSKSLLSALLAEMDKVGHVAKGSVAVYPQOAMTQND 720
 QY 721 SLRENILFGCOLLEPPYRSVIOACALLPDLILPSGDRTEIGEGVNLSSGQKORVSLAR 780
 DB 721 SLRENILFGCOLLEPPYRSVIOACALLPDLILPSGDRTEIGEGVNLSSGQKORVSLAR 780
 QY 781 AVVSNADIYLEDPLSAVDAHVGHIFENVIQPGMKMLKNKRIIIVTSMSTLPQVDYITV 840
 DB 781 AVVSNADIYLEDPLSAVDAHVGHIFENVIQPGMKMLKNKRIIIVTSMSTLPQVDYITV 840
 QY 841 MSGGKISEMGSYQELRLARDGAFAEFLRTYASTEOBDAEENGVTGSGPGKEAQMENG 900
 DB 841 MSGGKISEMGSYQELRLARDGAFAEFLRTYASTEOBDAEENGVTGSGPGKEAQMENG 900
 QY 901 LVTDLSAGKOLQRLSSSSSSYSGDISRRHNSYAELOKAEAKKEFTWKLMEADKAQTGVKL 960
 DB 901 LVTDLSAGKOLQRLSSSSSSYSGDISRRHNSYAELOKAEAKKEFTWKLMEADKAQTGVKL 960
 QY 961 SVYWDYMKATGLFTSPFLSTFLFMCNHNYSALASNWLSLMTDDPIVNGQOEHTRKRLSVYG 1020
 DB 961 SVYWDYMKATGLFTSPFLSTFLFMCNHNYSALASNWLSLMTDDPIVNGQOEHTRKRLSVYG 1020
 QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGULVNRFSKEL 1080
 DB 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGULVNRFSKEL 1080
 QY 1081 DTVDMSIPEVVKMFGSLFNVIAGACIVILLATPIAIIIPPLGIYEFVQRFYVASSROL 1140
 DB 1081 DTVDMSIPEVVKMFGSLFNVIAGACIVILLATPIAIIIPPLGIYEFVQRFYVASSROL 1140
 QY 1141 KRLESVRSPPYSHFNFTLGVSVIRAFEEBERFIHOSDLKVDENQRAYPSIYANWMLA 1200
 DB 1141 KRLESVRSPPYSHFNFTLGVSVIRAFEEBERFIHOSDLKVDENQRAYPSIYANWMLA 1200
 QY 1201 VRLIECVNGCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTVYLNWLVMSSEMETNIVA 1260
 DB 1201 VRLIECVNGCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTVYLNWLVMSSEMETNIVA 1260
 QY 1261 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNRYCLARYRDLDPVLRHINVTNGG 1320
 DB 1261 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNRYCLARYRDLDPVLRHINVTNGG 1320
 QY 1320 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNRYCLARYRDLDPVLRHINVTNGG 1320
 DB 1320 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNRYCLARYRDLDPVLRHINVTNGG 1320

QY 1321 EKVIGVERTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 DB 1321 EKVIGVERTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 QY 1381 SGSLRKNLDPFSQYSDEEWTSLSLAHLKDFVSALPKLHDECAEGGENISVQGRQIVCL 1440
 DB 1381 SGSLRKNLDPFSQYSDEEWTSLSLAHLKDFVSALPKLHDECAEGGENISVQGRQIVCL 1440
 QY 1441 ABALLRRTKILIVDEATPAVDLETPDDLIOSTIRFOFEDCTVLTAAHLNMTMDYTRVYL 1500
 DB 1441 ABALLRRTKILIVDEATPAVDLETPDDLIOSTIRFOFEDCTVLTAAHLNMTMDYTRVYL 1500
 QY 1501 DKGEIOEYGAAPSDLLQORGLFYSMAXDAGLV 1531
 DB 1501 DKGEIOEYGAAPSDLLQORGLFYSMAXDAGLV 1531
 RESULT 9
 ID AAY78873 standard; protein; 1531 AA.
 XX AAY78873;
 AC AAY78873;
 XX 19-MAY-2000 (first entry)
 DT 19-MAY-2000 (first entry)
 XX
 DE Multidrug resistance protein (MRP) natural variant amino acid sequence.
 XX
 KM Multidrug resistance protein; MRP; human; anthracycline; Vinca alkaloid;
 KM epipodophyllotoxin; cancer; leukaemia.
 XX Homo sapiens.
 OS
 XX US6025473-A.
 PN 15-FEB-2000.
 PD 15-FEB-2000.
 XX
 PF 05-JUN-1995; 95US-00461384.
 XX
 PR 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX
 PA (TOOH) UNIV QUEBENS KINGSTON.
 XX
 PI Cole SPC, Deeley RG;
 PI WPI; 2000-181838/16.
 XX N-PSDB; AA290193.
 DR
 XX
 PT Isolated protein conferring multidrug resistance, to at least two drugs
 PT selected from anthracyclines, epipodophyllotoxins and Vinca alkaloids, on
 PT a drug sensitive mammalian cell.
 XX
 PS Claim 10; Col 79-88; 78pp; English.
 XX
 CC This sequence represents a human multidrug resistance protein (MRP)
 CC natural variant amino acid sequence. The human MRP confers multidrug
 CC resistance, including resistance to at least two drugs selected from
 CC anthracyclines, epipodophyllotoxins and Vinca alkaloids, on a drug
 CC sensitive mammalian cell, when the protein is expressed in the cell. The
 CC multidrug resistance is not substantially reversed by chemosensitizers
 CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP
 CC protein sequence can be used to generate antibodies against MRP. The MRP
 CC protein and nucleotide sequences can be used in compositions which are
 CC used to treat patients with tumours displaying multidrug resistance. The
 CC compositions and methods of the invention can be used particularly to
 CC treat breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,
 CC thymomas, neuroblastomas, and lung cancer. Antibodies directed against
 CC MRP can be used to inhibit the multidrug resistance of a multidrug
 CC resistant cell
 XX
 SO Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MALRPGSADSDSLMDNNVTMTNSPDKCFQNTVTVWVPCFLTAMCFPFYLYSRH	60
Db	1	MALRPGSADSDSLMDNNVTMTNSPDKCFQNTVTVWVPCFLTAMCFPFYLYSRH	60
Qy	61	DRGIQMTPLNKKTALGFLIMIVCMADLFYSPWERSRGIFLAPVFLSPITLGLITTLA	120
Db	61	DRGIQMTPLNKKTALGFLIMIVCMADLFYSPWERSRGIFLAPVFLSPITLGLITTLA	120
Qy	121	TFLLQLEBRKGVQSSGIMLTFWLVALVICALILRSKIMTALKBDAQVDFPDITFYVFS	180
Db	121	TFLLQLEBRKGVQSSGIMLTFWLVALVICALILRSKIMTALKBDAQVDFPDITFYVFS	180
Qy	181	LLLLQVLVSCFSDSPLESETIHDPNCPESASLSTITTFWMTTGLIVRGYRQPLESSD	240
Db	181	LLLLQVLVSCFSDSPLESETIHDPNCPESASLSTITTFWMTTGLIVRGYRQPLESSD	240
Qy	241	LMSLNKEDTSQOVVPLVKNMKKECAKTRKQPVKVVSCKDPAQPKSSKYDANEVVAL	300
Db	241	LMSLNKEDTSQOVVPLVKNMKKECAKTRKQPVKVVSCKDPAQPKSSKYDANEVVAL	300
Qy	301	IVKSPQKEMNPSLFKVLYKTGYPFLMSPFKKAHDLMMFSGPQILKLIFVNDTKAPD	360
Db	301	IVKSPQKEMNPSLFKVLYKTGYPFLMSPFKKAHDLMMFSGPQILKLIFVNDTKAPD	360
Qy	361	WQGFYTYLLFVTACLOTLVNHQYHICFVSGMRKTAIVIGAVYKALVITNSAKSSTV	420
Db	361	WQGFYTYLLFVTACLOTLVNHQYHICFVSGMRKTAIVIGAVYKALVITNSAKSSTV	420
Qy	421	GEIVNLMSVDKORFMDLATYINMISAPLOVILAYLMLNGPVLGAVAVMVLVAVN	480
Db	421	GEIVNLMSVDKORFMDLATYINMISAPLOVILAYLMLNGPVLGAVAVMVLVAVN	480
Qy	481	AVMAKTKTYOVANHKSKONRIKLMBJLNGIKVLKUYAMELAFKDYALIROEELKYLK	540
Db	481	AVMAKTKTYOVANHKSKONRIKLMBJLNGIKVLKUYAMELAFKDYALIROEELKYLK	540
Qy	541	KSAYLSANGTPTWCTPFLVALCTPAVYVITDENNILDQAQAFVSLAFNLIRPLNTLP	600
Db	541	KSAYLSANGTPTWCTPFLVALCTPAVYVITDENNILDQAQAFVSLAFNLIRPLNTLP	600
Qy	601	MVISIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGGNSITVRNATFTWASDDEPT	660
Db	601	MVISIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGGNSITVRNATFTWASDDEPT	660
Qy	661	LNGITFSLPEGALVAVVGVGGKSSLLSALLAEMDKVEGHVAIKGSVAAYVPOQAMIOND	720
Db	661	LNGITFSLPEGALVAVVGVGGKSSLLSALLAEMDKVEGHVAIKGSVAAYVPOQAMIOND	720
Qy	721	SLRENIILFGCCLEBYRYSVIOACALLPDLLEILBPGDRTEIGEKGVNLSGGOKQVSLAR	780
Db	721	SLRENIILFGCCLEBYRYSVIOACALLPDLLEILBPGDRTEIGEKGVNLSGGOKQVSLAR	780
Qy	781	AVYSNADLYLPDDPLSAVDAAVGHKHFENVIGPKQMLKNKRIILVTHSMYSYPOVDVIV	840
Db	781	AVYSNADLYLPDDPLSAVDAAVGHKHFENVIGPKQMLKNKRIILVTHSMYSYPOVDVIV	840
Qy	841	MSGKISMGYSYOELLARDGAFAEFLRTYASTEOBDAEENGVTGVSQPKAKOMENGM	900
Db	841	MSGKISMGYSYOELLARDGAFAEFLRTYASTEOBDAEENGVTGVSQPKAKOMENGM	900
Qy	901	LVTDSAGQOLQRQLSSSSSSYSGDISRHNSTALQKAEKKETWKLMEADPAQTGOVYL	960
Db	901	LVTDSAGQOLQRQLSSSSSSYSGDISRHNSTALQKAEKKETWKLMEADPAQTGOVYL	960
Qy	961	SVYMDYMAKIGLFLISFLFMCHVSLASNYMLSLMTDDPIYNGTOEHTKVALSYVG	1020
Db	961	SVYMDYMAKIGLFLISFLFMCHVSLASNYMLSLMTDDPIYNGTOEHTKVALSYVG	1020

Qy	1021	ALGISGIAVFGSMANVSGIGILASRCLHYDLDLSILRSPMSFERTPSGNLVNRFSEKL	1080
Db	1021	ALGISGIAVFGSMANVSGIGILASRCLHYDLDLSILRSPMSFERTPSGNLVNRFSEKL	1080
Qy	1081	DTVDSMIPEVYKMGMSLFENVIGACIYILLATPIAIIIPPLGIYFFVQRFYVASSROL	1140
Db	1081	DTVDSMIPEVYKMGMSLFENVIGACIYILLATPIAIIIPPLGIYFFVQRFYVASSROL	1140
Qy	1141	KRLSVSRSPVYSHNFTLLGVSVIRAFEEQEREIHOSDLKVDENOKAYYPSIVANRWLA	1200
Db	1141	KRLSVSRSPVYSHNFTLLGVSVIRAFEEQEREIHOSDLKVDENOKAYYPSIVANRWLA	1200
Qy	1201	VRLCEVGNCTVLPALPAVVISRHSLSAGLVLSYSYLOVTTYANMLVRMSSEMETNIVA	1260
Db	1201	VRLCEVGNCTVLPALPAVVISRHSLSAGLVLSYSYLOVTTYANMLVRMSSEMETNIVA	1260
Qy	1261	VERLKEYSETEKEAPMOIOETAPPSMPOYGRVFRNYCLRYREBDLPVLHINVTINGG	1320
Db	1261	VERLKEYSETEKEAPMOIOETAPPSMPOYGRVFRNYCLRYREBDLPVLHINVTINGG	1320
Qy	1321	EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIPODPVLF	1380
Db	1321	EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIPODPVLF	1380
Qy	1381	SGSLRNLDPFSQYSDSEVWTSLELAHLKDFVSLPKDLHECABGGENLSVGQROLVCL	1440
Db	1381	SGSLRNLDPFSQYSDSEVWTSLELAHLKDFVSLPKDLHECABGGENLSVGQROLVCL	1440
Qy	1441	ARALLRKTILVLEATPAVVDLETDDLIQSTIRFOFEDCTVLTAAHRLNTIMDYTRVVL	1500
Db	1441	ARALLRKTILVLEATPAVVDLETDDLIQSTIRFOFEDCTVLTAAHRLNTIMDYTRVVL	1500
Qy	1501	DKGEIOEYGAPSDLLQQRGLFYSNAKDAGLV	1531
Db	1501	DKGEIOEYGAPSDLLQQRGLFYSNAKDAGLV	1531

RESULT 10		
ABG61810	ID	ABG61810 standard; protein; 1531 AA.
XX	XX	ABG61810;
AC	AC	15-AUG-2002 (first entry)
DT	DT	Prostate cancer-associated protein #1.
DE	DE	Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
KM	KM	Mammalia.
OS	OS	WO200230268-A2.
PN	PN	18-APR-2002.
XX	XX	12-OCT-2001; 2001WO-US032045.
PD	PD	13-OCT-2000; 2000US-00687576.
XX	XX	08-DEC-2000; 2000US-00733288.
PR	PR	08-DEC-2000; 2000US-00733742.
PR	PR	24-JAN-2001; 2001US-0263957P.
PR	PR	16-MAR-2001; 2001US-0276791P.
PR	PR	16-MAR-2001; 2001US-0276888P.
PR	PR	06-APR-2001; 2001US-0281922P.
PR	PR	24-APR-2001; 2001US-0286214P.
PR	PR	30-APR-2001; 2001US-00847046.
PR	PR	04-MAY-2001; 2001US-0288589P.
XX	XX	(BOSB-) EOS BIOTECHNOLOGY INC.
PA	PA	Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
PI	PI	WPI, 2002-471335/50.
XX	XX	
DR	DR	

DR N-PSDB; ABK92125.
XX Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
XX Claim 27; Page 309; 436pp; English.
XX
XX The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC associating a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridize to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
SQ Sequence 1531 AA;
Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGPSADSDLDLMMNTMTNTSNPTFCQNTLVLPWCYTLACFPFYLYLSRH 60
DB 1 MALRGPSADSDLDLMMNTMTNTSNPTFCQNTLVLPWCYTLACFPFYLYLSRH 60
QY 61 DRGTYIQTPLNKTALGFLIMTYCMADLFYSFWERSGFIAPVFLVPTLIGITLLA 120
DB 61 DRGTYIQTPLNKTALGFLIMTYCMADLFYSFWERSGFIAPVFLVPTLIGITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVALVCAALIRSKIMTALKEDAOVDLFFDITFYVFS 180
DB 121 TFLIQLERRKGVSSGIMLTFWLVALVCAALIRSKIMTALKEDAOVDLFFDITFYVFS 180
QY 181 LLLIQLVLSGSDSPFSETIHDNPPCESASFLSRTFWMITGLIVRGYRPLGSD 240
DB 181 LLLIQLVLSGSDSPFSETIHDNPPCESASFLSRTFWMITGLIVRGYRPLGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKYVSSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKYVSSKDPAPKESKVDANEVEAL 300
QY 301 IVKSPOKEMNPDLFKVLYKTGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTRAPD 360
DB 301 IVKSPOKEMNPDLFKVLYKTGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTRAPD 360
QY 361 WQGFYVTVLLFVTAQCLQTVLHOFPHICFVSGMRITKRAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYVTVLLFVTAQCLQTVLHOFPHICFVSGMRITKRAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQRFMDLATTYINMISAPLQVITLALYLLMLNGPSVLAVWMLPVN 480
DB 421 GEIYNLMSVDAQRFMDLATTYINMISAPLQVITLALYLLMLNGPSVLAVWMLPVN 480
QY 481 AVNMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLYKLYAMELAKDRLVALIROLELVK 540
DB 481 AVNMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLYKLYAMELAKDRLVALIROLELVK 540
QY 541 KSAVLAVGTFTWCTPFIVALCTFAVYVITDENNNIIDAQAFSLFNLFRPLNLP 600
DB 541 KSAVLAVGTFTWCTPFIVALCTFAVYVITDENNNIIDAQAFSLFNLFRPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGTNSITVENATFTWASDPPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGTNSITVENATFTWASDPPT 660

DB 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGTNSITVENATFTWASDPPT 660
QY 661 LNTGTTTIPGALVAVVGVCCGSSLLSALLMMDVBEHVAAKGSVAVYPOQAWIOND 720
DB 661 LNTGTTTIPGALVAVVGVCCGSSLLSALLMMDVBEHVAAKGSVAVYPOQAWIOND 720
QY 721 SLRENIIFGCOLBEPYRYSYIQAACALPDDEILPDSGRTEIGEGVNLSSGQKQVSLAR 780
DB 721 SLRENIIFGCOLBEPYRYSYIQAACALPDDEILPDSGRTEIGEGVNLSSGQKQVSLAR 780
QY 781 AVYSNADIVLPDDELSAVDAHVGKHIPENYIGPKMLKNKTRILVTHSMSYLPQVDYIIV 840
DB 781 AVYSNADIVLPDDELSAVDAHVGKHIPENYIGPKMLKNKTRILVTHSMSYLPQVDYIIV 840
QY 841 MSGGKISEMSYQELARPDGAFAPFLTYASTEOBOAENGVTVSGPGKEAQMENG 900
DB 841 MSGGKISEMSYQELARPDGAFAPFLTYASTEOBOAENGVTVSGPGKEAQMENG 900
QY 901 LVTDASAKOLQROULSSSSYSYSDISRHNSTAELOKAEKKEETWKLMEADKAQTGVKL 960
DB 901 LVTDASAKOLQROULSSSSYSYSDISRHNSTAELOKAEKKEETWKLMEADKAQTGVKL 960
QY 961 SYVWDYMKALGLFISFLSIFLFCMNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SYVWDYMKALGLFISFLSIFLFCMNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQGIAPFGISMAVSIQIILASRCLAYDLHSILRSPMSFPERPSCNLVNFSEKEL 1080
DB 1021 ALGISQGIAPFGISMAVSIQIILASRCLAYDLHSILRSPMSFPERPSCNLVNFSEKEL 1080
QY 1081 DTWDSMIPEVYIKPMGSLFNVIQACIVITLAPPAIIPPLGITYFVORFYVASSROL 1140
DB 1081 DTWDSMIPEVYIKPMGSLFNVIQACIVITLAPPAIIPPLGITYFVORFYVASSROL 1140
QY 1141 KRLESYSRSPVYSHFNETLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANWMLA 1200
DB 1141 KRLESYSRSPVYSHFNETLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANWMLA 1200
QY 1201 VRLBVCNGLVLPALFAVISRHSLSAGLVGSYSVLOVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLBVCNGLVLPALFAVISRHSLSAGLVGSYSVLOVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMIOETAPPSMPOVGRVFRNYCLARYEDLDPVLRHINVTING 1320
DB 1261 VERLKEYSETEKAPMIOETAPPSMPOVGRVFRNYCLARYEDLDPVLRHINVTING 1320
QY 1321 EKVGIIVGTAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIIVGTAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SSSIIRNMLDPPSOYSDEEVTSLLELAHLKDPVSALPKLDHECAEGGENLSVGQROLVCL 1440
DB 1381 SSSIIRNMLDPPSOYSDEEVTSLLELAHLKDPVSALPKLDHECAEGGENLSVGQROLVCL 1440
QY 1441 ARALLRRTKTLVDEATAAADLETDLLOSTIRTOPEDDCVLTATAHLANTIMYTRIVL 1500
DB 1441 ARALLRRTKTLVDEATAAADLETDLLOSTIRTOPEDDCVLTATAHLANTIMYTRIVL 1500
QY 1501 DKGEIOEYGAAPSDLLQORGLFYSMAKDAGIV 1531
DB 1501 DKGEIOEYGAAPSDLLQORGLFYSMAKDAGIV 1531
RESULT 11
ID ABM35012 standard; protein; 1531 AA.
XX
XX ABM35012;
XX
XX 08-OCT-2003 (first entry)
XX
XX Cancer based on CYP3A5 related protein SEQ ID NO:678.
XX

Db 1441 ARAIRKTKILVLEATVAVDLEFDDLIQSTIRQFEDCTVLTIAHRLNTIMDTYRVL 1500
QY 1501 DKGEIOEGAPSDLLQQRGLFYMAKQAGLV 1531
Db 1501 DKGEIOEGAPSDLLQQRGLFYMAKQAGLV 1531

RESULT 12
ADB20865
ID ADB20865 standard; protein, 1531 AA.
XX
AC ADB20865;
XX
DT 20-NOV-2003 (first entry)
XX
DE MRP1 based cancer related protein SEQ ID NO:678.
XX
KM irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KM lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KM variant allele; multidrug resistance protein 1; MRP1; cytosolic.
OS Unidentified.
XX
PN WO2003013533-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002WO-EP008200.
XX
PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX WPI; 2003-354397/33.
DR
PT Use of irinotecan or its derivative for preparation of a pharmaceutical
PT composition for treating cancer in a subject having a genome with a
PT variant allele comprising a multidrug resistance protein 1
PT polynucleotide.
XX
PS Disclosure; SEQ ID NO 678; 100pp; English.
XX
XX The present invention describes a method for the use of irinotecan (I) or
XX its derivative for the preparation of a pharmaceutical composition for
XX treating colorectal, cervical, gastric, lung, ovarian or pancreatic
XX cancer, or malignant glioma in a subject having a genome with a variant
XX allele which comprises a multidrug resistance protein 1 (MRP1)
XX polynucleotide (II). (I) has cytostatic activity. (I) or its derivative
XX can be used for the preparation of a pharmaceutical composition for
XX treating colorectal, cervical, gastric, lung, ovarian or pancreatic
XX cancer, or malignant glioma in a subject, where the subject is a human
XX (preferably African or Asian) or a mouse. The present sequence represents
XX a sequence which is used in the exemplification of the present invention.
XX

Sequence 1531 AA;
Query Match 79.4%; Score 7860; DB 6; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLMDMNTVNTSNDFPKCFONTLVVWPCFYLMACEFFYLYSRH 60
Db 1 MALRGFCSDGSDPLMDMNTVNTSNDFPKCFONTLVVWPCFYLMACEFFYLYSRH 60
QY 61 DRGICQMTPLNKTALGFLMLIVCADLRYSFERSRGIFLAPVFLVSPLLGITTLLA 120
Db 61 DRGICQMTPLNKTALGFLMLIVCADLRYSFERSRGIFLAPVFLVSPLLGITTLLA 120
QY 121 TPLIQLERRRGVSGSGLMFLFWALVCAALILRSKIMTLAKEDAQVDLFRDITFYVYFS 180

Db 121 TPLIQLERRRGVSGSGLMFLFWALVCAALILRSKIMTLAKEDAQVDLFRDITFYVYFS 180
QY 181 LLLIQLVLSGFSDBSLFSETIHDNPNCPSSSASFSLRITFWWTGLIVRGYRPLBGSD 240
Db 181 LLLIQLVLSGFSDBSLFSETIHDNPNCPSSSASFSLRITFWWTGLIVRGYRPLBGSD 240
QY 241 LMSLNKEDTSBOVVPVLYKMKKCEACTRRQOPVNVVSSKDPAPKSSKYDAEVEAL 300
Db 241 LMSLNKEDTSBOVVPVLYKMKKCEACTRRQOPVNVVSSKDPAPKSSKYDAEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLYKTFGPYFLMSFFFKAIHDMFSGPOLIKLILKFVNDTAPD 360
Db 301 IVKSPQKEMNPSLFKVLYKTFGPYFLMSFFFKAIHDMFSGPOLIKLILKFVNDTAPD 360
QY 361 WQGYFTVLLFVTAQCTLVLAHQYFHLCFVSGMKRIKTAIVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTAQCTLVLAHQYFHLCFVSGMKRIKTAIVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRMDLATYINMWSAPLOVILALYLMNLGFSVLGAVAVVLMVFN 480
Db 421 GEIVNLMSVDAQRMDLATYINMWSAPLOVILALYLMNLGFSVLGAVAVVLMVFN 480
QY 481 AVNAAKTKTYQVAHMSKDNRIKLMNEILNGIKVLYLAMELAFKDYALROBELKYLK 540
Db 481 AVNAAKTKTYQVAHMSKDNRIKLMNEILNGIKVLYLAMELAFKDYALROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVACTFAVYVTTDENNNLDAQTAFAVSLAFNLRPPLNLP 600
Db 541 KSAVLSAVGTFTWCTPFLVACTFAVYVTTDENNNLDAQTAFAVSLAFNLRPPLNLP 600
QY 541 KSAVLSAVGTFTWCTPFLVACTFAVYVTTDENNNLDAQTAFAVSLAFNLRPPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVQDGGGNTSTVENAFTWARSPPPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVQDGGGNTSTVENAFTWARSPPPT 660
QY 661 LMGITFSIPGALVAVVGQVCGGKSLSLALMEMDVEGHVMIKGSVAVPOQAMQND 720
Db 661 LMGITFSIPGALVAVVGQVCGGKSLSLALMEMDVEGHVMIKGSVAVPOQAMQND 720
QY 721 SURENLIIFGQLEBPYRYSVIOACALLPDLIELPSGDRTEIGEGVNLASGQKORVSLAR 780
Db 721 SURENLIIFGQLEBPYRYSVIOACALLPDLIELPSGDRTEIGEGVNLASGQKORVSLAR 780
QY 781 AVYSNADYIFPDPLSANDAVHGHIFENYIGRKMLKNTKTRILLVTHSMGYLPOVDYIIV 840
Db 781 AVYSNADYIFPDPLSANDAVHGHIFENYIGRKMLKNTKTRILLVTHSMGYLPOVDYIIV 840
QY 841 MSGGKISEMGSYOELLARDGAFAPLRTVASTOEOBDAEENGVTGVSQPKKAKOMENGM 900
Db 841 MSGGKISEMGSYOELLARDGAFAPLRTVASTOEOBDAEENGVTGVSQPKKAKOMENGM 900
QY 901 LVTDSAGKQLOROLSSSSSYSGDISRHNNSTAELOKAEKKEBTWKLMEADKAQTQGVKL 960
Db 901 LVTDSAGKQLOROLSSSSSYSGDISRHNNSTAELOKAEKKEBTWKLMEADKAQTQGVKL 960
QY 961 SVYWDYMKALGFLISFLSIFLFWCNHVSALASVWYLSLWTDPIVNGTOEHTYVRLSVYG 1020
Db 961 SVYWDYMKALGFLISFLSIFLFWCNHVSALASVWYLSLWTDPIVNGTOEHTYVRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSTGILLASRCILHVDLHSLISRMSFFERTPSGNLVNRSKEL 1080
Db 1021 ALGISOGIAVFGYSMAVSTGILLASRCILHVDLHSLISRMSFFERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPEVIKMFMSGLFNVIGACIVILATPIAIIIPPLGLIYFFVORFYVASSROL 1140
Db 1081 DTVDSMIPEVIKMFMSGLFNVIGACIVILATPIAIIIPPLGLIYFFVORFYVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNMTLLIGSVIVRAFEQOEFRHOSDLKVDENQKAYYPSIVANRWLA 1200
Db 1141 KRLSVSRSPVYSHFNMTLLIGSVIVRAFEQOEFRHOSDLKVDENQKAYYPSIVANRWLA 1200
QY 1201 VRLCYGNCIVFAALFAVYSRSLSAGLVGSLVSVSLQVTTVLMNLVRRSSMETNIVA 1260

Db 1201 VRLECVGNCIVLPAALFAVVISRHSLSAGLVGLSVSYSLQVTTVLTNMLVMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKEA PMQIOETAPSPSPQVGRVEFRNYCLRYREDLDFVLHINVTINGG 1320
Db 1261 VERLKEVSETEKEA PMQIOETAPSPSPQVGRVEFRNYCLRYREDLDFVLHINVTINGG 1320
Qy 1321 EKVGIVGRTGAKSSLTGLFRINESAGEI11IDGINIAK1GLHDLRKIT111PODPVL 1380
Db 1321 EKVGIVGRTGAKSSLTGLFRINESAGEI11IDGINIAK1GLHDLRKIT111PODPVL 1380
Qy 1381 SGSLRMNLDPRFSQVSDSEVMTSLAHKDPFVSLPDLHDHRCAGENLSVGORQIVCL 1440
Db 1381 SGSLRMNLDPRFSQVSDSEVMTSLAHKDPFVSLPDLHDHRCAGENLSVGORQIVCL 1440
Qy 1441 ARALLRKTIIVLDEATAVAVLETDLLIOSTIRIQFEDCTVLT1AHRNTIMDYTRVIVL 1500
Db 1441 ARALLRKTIIVLDEATAVAVLETDLLIOSTIRIQFEDCTVLT1AHRNTIMDYTRVIVL 1500
Qy 1501 DKGEIQEYGAPSDLIQQRGLFFYSMAKDAGLV 1531
Db 1501 DKGEIQEYGAPSDLIQQRGLFFYSMAKDAGLV 1531

RESULT 13

ADB87954 standard; protein; 1531 AA.

ADB87954;

04-DEC-2003 (first entry)

Human UGT1A1 protein sequence SEQ ID NO:678.

irrinotecan; cancer; UGT1A1; cytostatic; topoisomerase I inhibitor;
colorectal cancer; cervical cancer; gastric cancer; lung cancer;
ovarian cancer; pancreatic cancer; malignant glioma;
uridine diphosphate glycosyltransferase1 member A1.

Homo sapiens.

WO2003013536-A2.

20-FEB-2003.

23-JUL-2002; 2002MO-EP008217.

23-JUL-2001; 2001EP-00117608.

24-MAY-2002; 2002EP-00011710.

(EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

Heinrich G, Kerb R;

WPI; 2003-289896/28.

Use of irinotecan to treat cancer patient by determining if patient has
variant alleles of UGT1A1 gene, administering increased/decreased amounts
of irinotecan based on increased/decreased levels of UGT1A1 gene product.

Disclosure: SEQ ID NO 678; 107bp; English.

The invention relates to the novel use of irinotecan to treat a patient
suffering from cancer. This involves determining if the patient has one
or more variant alleles of the UGT1A1 gene, and if the patient has one
or more of such variant alleles, irinotecan is administered in an increased
or decreased amount in comparison to the amount that is administered
without regard to the patient's alleles in the UGT1A1 gene. The invention
has cytostatic activity. A composition of the invention acts as a
topoisomerase I inhibitor. The method is useful for treating a patient,
an animal e.g. mouse or a human, preferably African or Asian, suffering
from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
pancreatic cancer or malignant glioma. The present sequence is udes in
the exemplification of the invention.

XX Sequence 1531 AA:

Query Match Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGCSADGSDPLMDMNTWNTSNPDFTKCFONTYLVVWPCFYLAACPFYLYLSRH 60
Db 1 MALRGCSADGSDPLMDMNTWNTSNPDFTKCFONTYLVVWPCFYLAACPFYLYLSRH 60
Qy 61 DRGYIOMTPNKTATAGELIMYICWADLFYSFMRSGIFLAIVPLVSPLLGITTLA 120
Db 61 DRGYIOMTPNKTATAGELIMYICWADLFYSFMRSGIFLAIVPLVSPLLGITTLA 120
Qy 121 TFLIQERRKGVSSGIMLTFEVLVALYCALAILSKIMTALKEDAQVDLFEIDTFYVYS 180
Db 121 TFLIQERRKGVSSGIMLTFEVLVALYCALAILSKIMTALKEDAQVDLFEIDTFYVYS 180
Qy 181 LLLIQVLVSCFSDRSPLESETIHDNPPCESSASFLSRTFTWITGLIVRGYROPLEBSD 240
Db 181 LLLIQVLVSCFSDRSPLESETIHDNPPCESSASFLSRTFTWITGLIVRGYROPLEBSD 240
Qy 241 LMSINKEDTSROVVPVLYKMKKECAKTRKQPVKVYVSSKQPAQKSSKYDANEVEAL 300
Db 241 LMSINKEDTSROVVPVLYKMKKECAKTRKQPVKVYVSSKQPAQKSSKYDANEVEAL 300
Qy 301 IVKSPKEMNPSLFEVLYKTFGPFYFLMSFFPKALHDLMMFSGPOLKLLIKFVNDTKAPD 360
Db 301 IVKSPKEMNPSLFEVLYKTFGPFYFLMSFFPKALHDLMMFSGPOLKLLIKFVNDTKAPD 360
Qy 361 WQGFYFVTLFVTAQCLQTLVHQYFHI CFVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGFYFVTLFVTAQCLQTLVHQYFHI CFVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLSVVDQRMDLATYINMTWSAPLOVITLXVLMNLSGVLAVGVVMTLMVYN 480
Db 421 GEIVNLSVVDQRMDLATYINMTWSAPLOVITLXVLMNLSGVLAVGVVMTLMVYN 480
Qy 481 AVMAKKTQYVAAHMKSKDNRIKIMNELINGIKVLYKLYAMELAFKDYALIROBELKYLK 540
Db 481 AVMAKKTQYVAAHMKSKDNRIKIMNELINGIKVLYKLYAMELAFKDYALIROBELKYLK 540
Qy 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVTTIDENNILDAQAFVSLALEFNILRPPLNITL 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVTTIDENNILDAQAFVSLALEFNILRPPLNITL 600
Qy 601 MYISSIVASVSLKRLRIFLSHELEBPSIRRRVYKGGGINSITVRATFTWARSDDPT 660
Db 601 MYISSIVASVSLKRLRIFLSHELEBPSIRRRVYKGGGINSITVRATFTWARSDDPT 660
Qy 661 LMGTFPSIPEGALVAVVGVQVCGKSSLSALLAEMDKVEGHVAILKGSVAVYVPOQAMIOND 720
Db 661 LMGTFPSIPEGALVAVVGVQVCGKSSLSALLAEMDKVEGHVAILKGSVAVYVPOQAMIOND 720
Qy 721 SLRENILFCQLEBRYRSVIOACALLPDLILPSGDRTEIGERGVNLSGGOKORVSLAR 780
Db 721 SLRENILFCQLEBRYRSVIOACALLPDLILPSGDRTEIGERGVNLSGGOKORVSLAR 780
Qy 781 AVYSNADILYLPDDPLSAVDAAVGHKIPFNVI GPKKMLKMKRILVTHMSNLTLPQVDVIV 840
Db 781 AVYSNADILYLPDDPLSAVDAAVGHKIPFNVI GPKKMLKMKRILVTHMSNLTLPQVDVIV 840
Qy 841 MSGGKISMGVSOELIANDGAFAEFLRTYASTBOBDAEENGVTGVSPGKBAKOMENGM 900
Db 841 MSGGKISMGVSOELIANDGAFAEFLRTYASTBOBDAEENGVTGVSPGKBAKOMENGM 900
Qy 901 LVTSAGKOLOROLSSSSSSSGDISRRHNSYAELOKAKKEETWKLMEADKAGQGYKL 960
Db 901 LVTSAGKOLOROLSSSSSSSGDISRRHNSYAELOKAKKEETWKLMEADKAGQGYKL 960
Qy 961 SYVMDYKALIGLFISSLSIFPMCHVSAALSNVWLSLWTDPIVNGOERTKVLSTYVG 1020
Db 961 SYVMDYKALIGLFISSLSIFPMCHVSAALSNVWLSLWTDPIVNGOERTKVLSTYVG 1020

Db 961 SVYWDYMKAIIGLFIISFISIFLFCMCHVSAALASNYMLSLMTDPIVNGTQHEKTVRLSVYG 1020
QY 1021 ALGISOGIAVGYSMAYSIGIIASRCLHNDLHSIISSPMSFEERPPSGULVNRFSKEL 1080
Db 1021 ALGISOGIAVGYSMAYSIGIIASRCLHNDLHSIISSPMSFEERPPSGULVNRFSKEL 1080
QY 1081 DTVDMSIPEVYIKMGMGLFNVIIGACIYILLATPIAIIIPPLGIFYFVQGFYVASSROL 1140
Db 1081 DTVDMSIPEVYIKMGMGLFNVIIGACIYILLATPIAIIIPPLGIFYFVQGFYVASSROL 1140
QY 1141 KRLESVSRSPYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANMELA 1200
Db 1141 KRLESVSRSPYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANMELA 1200
QY 1201 VRLECVCNCIYLFALFAVVISRHSLSAGLVGLSVSYSLQVTTYANMLVNRMSSEMETNIVA 1260
Db 1201 VRLECVCNCIYLFALFAVVISRHSLSAGLVGLSVSYSLQVTTYANMLVNRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPPSWPOVGRVFRNYCLRYREDLDFVLRIHIVTNGG 1320
Db 1261 VERLKEYSETEKEAPMOIOETAPPSWPOVGRVFRNYCLRYREDLDFVLRIHIVTNGG 1320
QY 1321 EKVGIIVGRTAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIIVGRTAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRMLDPPSOYSDSEWNTSLBLAKDFVSLPDKLDHECAGEGNSVGOQOLVCL 1440
Db 1381 SGSLRMLDPPSOYSDSEWNTSLBLAKDFVSLPDKLDHECAGEGNSVGOQOLVCL 1440
QY 1441 ABALLRKTKLIVDEATAVADLETDDLIQSTIRTOFEDCTVLTARHLANTIMDTRYIVL 1500
Db 1441 ABALLRKTKLIVDEATAVADLETDDLIQSTIRTOFEDCTVLTARHLANTIMDTRYIVL 1500
QY 1501 DKGEIOEYGA PSDLQORGLFYSMAXDAGLV 1531
Db 1501 DKGEIOEYGA PSDLQORGLFYSMAXDAGLV 1531

RESULT 14
ADB96937 standard; protein; 1531 AA.
XX ADB96937;
AC ADB96937;
DT 04-DEC-2003 (first entry)
XX
DE Human MDR1 related protein sequence SEQ ID NO:678.
XX
KM irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KM lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KM multidrug resistance 1; MDR1; cytosolic; human; Cyp3A5; MRP1; MDR1;
KM TOP1.
XX
OS Homo sapiens.
XX
PN W02003013537-A2.
XX
PD 20-FEB-2003.
PF 23-JUL-2002; 2002WC-EP008218.
PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-268145/26.
PT New use of irinotecan for preparation of pharmaceutical compositions for treating cancer in subject having genome with variant allele comprising

PT multidrug resistance 1 polynucleotide.
XX
PS Disclosure; SEQ ID NO 678; 130pp; English.
XX
CC The invention relates to the novel use of irinotecan or its derivative
CC for the preparation of pharmaceutical compositions for treating
CC colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or
CC malignant glioma in a subject having a genome with a variant allele which
CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition
CC of the invention has cytotoxic activity. The invention is useful for the
CC preparation of pharmaceutical compositions for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject (preferably human, more preferably African or Asian)
CC or a mouse. The present sequence is used in the exemplification of the
CC invention.
XX
SQ Sequence 1531 AA;
Query Match 79.4%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFQNTLVWVPCFYLMACEPFYLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFQNTLVWVPCFYLMACEPFYLYLSRH 60
QY 61 DRGVIQMTPLNKTKTALGFLMIVCWADLFYSFWEERSRGIFLAPVPLVSPTLGITTLA 120
Db 61 DRGVIQMTPLNKTKTALGFLMIVCWADLFYSFWEERSRGIFLAPVPLVSPTLGITTLA 120
QY 121 TPLIQLERRRGVSSGIMLTFMLVALVCAALIRSKTMTLXDAQVDRDITFYVYFS 180
Db 121 TPLIQLERRRGVSSGIMLTFMLVALVCAALIRSKTMTLXDAQVDRDITFYVYFS 180
QY 181 LLLIQLVLSGFSRSPFLSETHIDPNCPESSASFLSRITFWMTGLIVRGYRPLEGSD 240
Db 181 LLLIQLVLSGFSRSPFLSETHIDPNCPESSASFLSRITFWMTGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKOPKVVYSSKDPQPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKOPKVVYSSKDPQPKSSKVDANEVEAL 300
QY 301 IVKSPQKEMNPFLFKYLKTRGYPFLMSFFFKAIHDLMSGQILKLKFPVNDTAPD 360
Db 301 IVKSPQKEMNPFLFKYLKTRGYPFLMSFFFKAIHDLMSGQILKLKFPVNDTAPD 360
QY 361 WQGYFTVLLFVTAACIQTIVLHOFYFHCFSYGRKIKTAVGAYRKALVITNSARKSSTV 420
Db 361 WQGYFTVLLFVTAACIQTIVLHOFYFHCFSYGRKIKTAVGAYRKALVITNSARKSSTV 420
QY 421 GEIVNIMSVDAGRFMDLATYINMISAPLOVILALYLMLNLPISVLGAVAVMLVAVN 480
Db 421 GEIVNIMSVDAGRFMDLATYINMISAPLOVILALYLMLNLPISVLGAVAVMLVAVN 480
QY 481 AVNMAKTKTYQVAHMSKONRITLAMEIINGIKVLKYLAMELAFKQVLAIRQBELKVLK 540
Db 481 AVNMAKTKTYQVAHMSKONRITLAMEIINGIKVLKYLAMELAFKQVLAIRQBELKVLK 540
QY 541 KSAVYLSAVGFTFWCTPFLVATCTPAVYVYIDENNILDAQTAVSLAFNIIPLPILP 600
Db 541 KSAVYLSAVGFTFWCTPFLVATCTPAVYVYIDENNILDAQTAVSLAFNIIPLPILP 600
QY 601 MTISSIVQASVSLKRLIFLSHEELPDSIERRPVDDGGTNSITVMAATFTYARSDPT 660
Db 601 MTISSIVQASVSLKRLIFLSHEELPDSIERRPVDDGGTNSITVMAATFTYARSDPT 660
QY 661 LINGITSIPGALVAVVGVGCGKSSLSLALAMDKVBSGVAIKSVAVVPOAMIQND 720
Db 661 LINGITSIPGALVAVVGVGCGKSSLSLALAMDKVBSGVAIKSVAVVPOAMIQND 720
QY 721 SIRENITLFGQLEPEYRSYVIOACALLPDLIELPSGDREITGEGVNLGGQKORVSLAR 780
Db 721 SIRENITLFGQLEPEYRSYVIOACALLPDLIELPSGDREITGEGVNLGGQKORVSLAR 780

QY 781 AVYSNADIIYLFDDPLSAVDHVGKIFENYIGPKMKLNKTRILIVTHSMSTLPQVDVITV 840
DB 781 AVYSNADIIYLFDDPLSAVDHVGKIFENYIGPKMKLNKTRILIVTHSMSTLPQVDVITV 840
QY 841 MSGGKISEMGSYOBELARDGAFARLFTYASTEOBODAEENGVCVSGPGKEAKOMEGM 900
DB 841 MSGGKISEMGSYOBELARDGAFARLFTYASTEOBODAEENGVCVSGPGKEAKOMEGM 900
QY 901 LVYDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAEKKEETVKLEADKAOYGKYL 960
DB 901 LVYDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAEKKEETVKLEADKAOYGKYL 960
QY 961 SVYDYMKAIGLFTISFLIFMCMHVSALASNTWLSMTDDPIVNGTOEHTKVALSVYG 1020
DB 961 SVYDYMKAIGLFTISFLIFMCMHVSALASNTWLSMTDDPIVNGTOEHTKVALSVYG 1020
QY 1021 ALGISOGIARVGSMAVSTIGIILASRCIHLNLSHILSPMSFPERTSGLVNRFSKEL 1080
DB 1021 ALGISOGIARVGSMAVSTIGIILASRCIHLNLSHILSPMSFPERTSGLVNRFSKEL 1080
QY 1081 DTVDSMIPEVIKMFMSLFENVIGACIVILATPIAIIIPPLGLIYFVQRFYVASSRQL 1140
DB 1081 DTVDSMIPEVIKMFMSLFENVIGACIVILATPIAIIIPPLGLIYFVQRFYVASSRQL 1140
QY 1141 KRLSVSRSPVYSHNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYSIVANRWLA 1200
QY 1201 VRLSCVNCICVLPALFPAVIRSHSLASGLVSGVSYQVTTYLMVLRMSSEMTNIVA 1260
DB 1201 VRLSCVNCICVLPALFPAVIRSHSLASGLVSGVSYQVTTYLMVLRMSSEMTNIVA 1260
QY 1261 VERLKEVSETEKEAPWQIOETAPRPSWPQVGVVEFRNYCLARYEDLDVLRHINVTINGG 1320
DB 1261 VERLKEVSETEKEAPWQIOETAPRPSWPQVGVVEFRNYCLARYEDLDVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSSLTGLFRINSAGEIIIDGINIAKIGLHDLRFKITIIPQDPVLF 1380
DB 1321 EKVGIVGRTGAGKSSLTGLFRINSAGEIIIDGINIAKIGLHDLRFKITIIPQDPVLF 1380
QY 1381 SGSLRMLNDPFSQVSDSEVMTSLAHLKDPYSALPDLDHHCAGGENLSVGORQVYCL 1440
DB 1381 SGSLRMLNDPFSQVSDSEVMTSLAHLKDPYSALPDLDHHCAGGENLSVGORQVYCL 1440
QY 1441 ARALIRKTIIVLDEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ARALIRKTIIVLDEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMKADAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMKADAGLV 1531
RESULT 15
ADB92128
ID ADB92128 standard; protein; 1531 AA.
XX ADB92128;
XX 04-DEC-2003 (first entry)
XX Human MDR1 related protein sequence SEQ ID NO:678.
XX
XX irinotecan; colorectal cancer; cervical cancer; gastric cancer;
XX lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
XX multidrug resistance 1; MDR1; cytosolic; human; UGT1A1; MRP1; TOP1.
XX Homo sapiens.
XX OS
XX PN WO2003013535-A2.
XX PD 20-FEB-2003.

XX 23-JUN-2002; 2002WO-EP008220.
PF 23-JUN-2001; 2001EP-00117608.
XX 24-MAY-2002; 2002EP-00011710.
PR (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX Heinrich G, Korb R;
XX WPI; 2003-342400/32.
XX
XX New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.
XX
XX Disclosure; SEQ ID NO 678; 104pp; English.
XX
XX The invention relates to a novel use of irinotecan or its derivative for
CC the preparation of a pharmaceutical composition for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject having a genome with a variant allele which comprises
CC a multidrug resistance 1 (MDR1) polynucleotide. A composition of the
CC invention has cytostatic activity. The present sequence is used in the
CC exemplification of the invention.
XX
XX Sequence 1531 AA;
SQ
Query Match 79.4%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFSGADGSDPLMDMNTVMTNTSNPDFTKCFQNTVYLWVPCFYLMACPFYFLYSRH 60
DB 1 MALRGFSGADGSDPLMDMNTVMTNTSNPDFTKCFQNTVYLWVPCFYLMACPFYFLYSRH 60
QY 61 DRGYIQMTPLNKTITAGLFLIMIVCMADLPFSFMRBSRGITFLAVPLVPLSLGTTTLA 120
DB 61 DRGYIQMTPLNKTITAGLFLIMIVCMADLPFSFMRBSRGITFLAVPLVPLSLGTTTLA 120
QY 121 TFLIQLEBRKGVSSGIMLTFMLVALVCAALILSKIMTALKEBQVDLPFDITFYVYS 180
DB 121 TFLIQLEBRKGVSSGIMLTFMLVALVCAALILSKIMTALKEBQVDLPFDITFYVYS 180
QY 181 LLLIQVLVSCFSDSPFSETIHDNPPCESSASFLSRTFMTITGLIVRGYROPLEGSD 240
DB 181 LLLIQVLVSCFSDSPFSETIHDNPPCESSASFLSRTFMTITGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKKMKKSCAKTRKQPVVVVSSKDPAPKESSKVDANEVEVAL 300
DB 241 LMSLNKEDTSEQVVPVLVKKMKKSCAKTRKQPVVVVSSKDPAPKESSKVDANEVEVAL 300
QY 301 IVKSPQKEMNPSLFRVLYKTGFPYFLMSFFPKAIDHLMFSGPOLIKLLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFRVLYKTGFPYFLMSFFPKAIDHLMFSGPOLIKLLIFVNDTKAPD 360
QY 361 WQGYFTYVLFVTAQLOTLVLAHQYFHLCEVSGMRKTAIVAGVYRKALVINTSARKSSTV 420
DB 361 WQGYFTYVLFVTAQLOTLVLAHQYFHLCEVSGMRKTAIVAGVYRKALVINTSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMTWSAPLVQYIALYLLMLMGPSVLGAVVMVLMVEVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMTWSAPLVQYIALYLLMLMGPSVLGAVVMVLMVEVN 480
QY 481 AVMMKTKTYOVAMHKSNDNR IKLMNELNLNGIYKLYAMELAFDKYALROBELKYLK 540
DB 481 AVMMKTKTYOVAMHKSNDNR IKLMNELNLNGIYKLYAMELAFDKYALROBELKYLK 540
QY 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVYTDENNILDAQTAFSALFNLRPLNITLP 600
DB 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVYTDENNILDAQTAFSALFNLRPLNITLP 600
QY 601 MVISSIVQSVSLKRLRIFLSHELEPDSIERRPVKGGTNSITVRNATFTWASDPPT 660

Db 601 MVISSIVQASVSLKRLRIFLSHBELEPDSIRRRPVKOGGTSITVRNATFTWARSDEPT 660
Qy 661 LNCITFSTPEGALVAVVQVCCGKSSLSALLAEMDKYGHVAIKGSVAIVPOQAMIOND 720
Db 661 LNCITFSTPEGALVAVVQVCCGKSSLSALLAEMDKYGHVAIKGSVAIVPOQAMIOND 720
Qy 721 SLRENIIFGCOLLEBPYRSVIQACALPDLLEIPSGDRTEIGEGKVNLSGGOKORVSLAR 780
Db 721 SLRENIIFGCOLLEBPYRSVIQACALPDLLEIPSGDRTEIGEGKVNLSGGOKORVSLAR 780
Qy 781 AVYSNADIYLPDDPLSAVDHVKHIFENVIGPKMLNKXTRILVTHSMSYLPOVDVILV 840
Db 781 AVYSNADIYLPDDPLSAVDHVKHIFENVIGPKMLNKXTRILVTHSMSYLPOVDVILV 840
Qy 841 MSGKISMSGSYQELIARDGAFAEFLRTYASTEOQDAENGVTGVSQPGKEAKQEMNGM 900
Db 841 MSGKISMSGSYQELIARDGAFAEFLRTYASTEOQDAENGVTGVSQPGKEAKQEMNGM 900
Qy 901 LVTSAGKQLOROLSSSSSGDISRHHNSTAELOKAEKKEETMKLEADKAOTGOVKL 960
Db 901 LVTSAGKQLOROLSSSSSGDISRHHNSTAELOKAEKKEETMKLEADKAOTGOVKL 960
Qy 961 SVYDYMKAIGLFTSFLSIFLMCNHVSALASNYWLSLMTDDPIVNGTOEHTKYRLSYG 1020
Db 961 SVYDYMKAIGLFTSFLSIFLMCNHVSALASNYWLSLMTDDPIVNGTOEHTKYRLSYG 1020
Qy 1021 ALGISOGIAVFGYMAVSIIGGILASRCLHVDLHLSIRSPMSFPERTSGNLVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGGILASRCLHVDLHLSIRSPMSFPERTSGNLVNRFSKEL 1080
Qy 1081 DTVUSMTPEVTKMFMGSLFNVIGACIVILLATPTAIIIPPLGLIYFFVQRFYVASSROL 1140
Db 1081 DTVUSMTPEVTKMFMGSLFNVIGACIVILLATPTAIIIPPLGLIYFFVQRFYVASSROL 1140
Qy 1141 KRLESVRSRPFYSHFNETLGVSVYRAFEEOERFIHOSDLKYDENOKAYYPSIVANRWLA 1200
Db 1141 KRLESVRSRPFYSHFNETLGVSVYRAFEEOERFIHOSDLKYDENOKAYYPSIVANRWLA 1200
Qy 1201 VRELCVGNCIYLPALPAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSEMETNIVA 1260
Db 1201 VRELCVGNCIYLPALPAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFEFRNYCLAYREDLDPVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFEFRNYCLAYREDLDPVLRHINVTINGG 1320
Qy 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKTIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKTIIPQDPVLF 1380
Qy 1381 SGLRNMNLDPPSQYSDEEWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQROLVCL 1440
Db 1381 SGLRNMNLDPPSQYSDEEWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQROLVCL 1440
Qy 1441 ARALLRKTILVDEATPAVLETFDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKTILVDEATPAVLETFDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Qy 1501 DKGEIOEYGA PSDLQOQGLFYMAKDAGLV 1531
Db 1501 DKGEIOEYGA PSDLQOQGLFYMAKDAGLV 1531

Search completed: December 15, 2005, 15:17:03
Job time : 186.366 secs

1941-1942 (1941-1942)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:57:04 ; Search time 37,8285 Seconds
(without alignments)
4901.320 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903

Sequence: 1 MALRGFCSADGSDPLMDMNV.....RSVAVAKXKPFISPDLS 1927

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 80:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7849	79.3	1531	1	DVHUR
2	4484.5	45.3	1527	1	JB0336
3	3587	36.2	1545	1	S71841
4	3507.5	35.4	1541	1	S71839
5	3289	33.2	1494	2	B89447
6	3225	32.6	1573	2	T21219
7	3195.5	32.3	1502	2	T42216
8	2838	28.7	1515	1	S51863
9	2732	27.6	1398	2	T20434
10	2548	25.7	1478	2	T38712
11	2384.5	24.1	1623	2	T01369
12	2356.5	23.8	1622	2	D86428
13	2326	23.5	1495	2	B86428
14	2294.5	23.3	1144	2	T27408
15	2263.5	22.9	1559	2	S64757
16	2263.5	22.6	1488	2	B86428
17	2238.5	22.6	1516	2	F84919
18	2230	22.5	1539	2	T48059
19	2144.5	21.7	1335	2	T00961
20	2144.5	21.7	1514	2	T52080
21	2140	21.6	1515	2	T52081
22	2131.5	21.5	1490	2	T47840
23	2129	21.5	1545	2	T46645
24	2123	21.4	1153	2	T26883
25	2109	21.3	1545	2	T42751
26	2104.5	21.3	1511	2	T42711
27	2098	21.2	1546	2	T42728
28	2069.5	20.9	1389	2	T47796
29	2062.5	20.8	1661	2	S64800

30	2037	20.6	1121	2	C87973	protein Y43F8C.12
31	2014.5	20.3	1582	2	A56248	sulfonylurea recep
32	2008	20.3	1592	2	S48933	probable transp
33	2007	20.3	390	2	S68403	inward rectifier p
34	1992	20.1	390	2	UC4689	inwardly rectifyin
35	1935	19.5	390	2	A57616	inward rectifier K
36	1932	19.5	390	2	UC7901	inwardly-rectifyin
37	1888	19.1	1477	2	S64616	YORI protein - Yea
38	1884.5	19.0	1037	2	T50518	ABC transporter-1i
39	1805	18.2	1146	2	F84487	Probable ABC trans
40	1722	17.4	946	1	UC5667	multidrug resistan
41	1693	17.1	1548	1	DVLNS	multidrug resistan
42	1687	17.0	1421	2	T34225	hypothetical prote
43	1606	16.2	1427	2	T20903	hypothetical prote
44	1575.5	15.9	1427	2	T39219	atp-binding caset
45	1558	15.7	1469	2	T50210	probable ABC trans

ALIGNMENTS

RESULT 1

DVHUR

multidrug resistance protein (cell line H69AR) - human

N:Alternate names: multidrug resistance-associated protein (MRP)

C:Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001

C:Accession: A44231; A37495

R:Coile, S.P.C.; Bhargava, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C., Science 258, 1650-1654, 1992

A:Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer

A:Reference number: A44231, MUID:93088080; PMID:1360704

A:Accession: A44231

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1613, 1-1613

A:Cross-references: UNIPARC:UPI00001746C; GB:105628; NID:91835658

A:Experimental source: small cell lung carcinoma cell line H69AR

A>Note: Sequence extracted from NCBI backbone (NCBI:119851); this sequence has been corrected.

A:Title: Multidrug resistance-associated protein: sequence correction.

A:Reference number: A37495; MUID:93262415; PMID:8098549

A:Accession: A37495

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-60 <CO2>

A:Cross-references: UNIPARC:UPI00001746C; GB:105628; NID:91835658

A>Note: sequence extracted from NCBI backbone (NCBI:131929)

C:GeneticB:

A:Gene: GDB:MRP

A:Cross-references: GDB:136335; OMIM:158343

A:Map position: 16p13.1-16p13.1

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C:Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmembrane

F:661-844/Domain: ATP-binding cassette homology <ABC1>

F:678-685/Region: nucleotide-binding motif A (P-loop)

F:110-1503/Domain: ATP-binding cassette homology <ABC2>

F:1327-1334/Region: nucleotide-binding motif A (P-loop)

F:1450-1454/Region: nucleotide-binding motif B

Query Match 79.3%; Score 7849; DB 1; Length 1531;

Best local similarity 99.9%; Pred. No. 0;

Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	MALRGFCSADGSDPLMDMNVNTNTNSNDFPKCFONTLVWVPCFYLMACPFYLYSRH	60
DB	1	MALRGFCSADGSDPLMDMNVNTNTNSNDFPKCFONTLVWVPCFYLMACPFYLYSRH	60
QY	61	DRGVQMTPLNKTALGFLMTVCVADLFYFWMERSRGIFLAPVFLVSPTLGITTLA	120
DB	61	DRGVQMTPLNKTALGFLMTVCVADLFYFWMERSRGIFLAPVFLVSPTLGITTLA	120

QY 121 TFLIQLERRKGVSSGIMLTFLMVALYCALAILRSKIMTALKEDAQVDLFDITFYFYS 180
 Db 121 TFLIQLERRKGVSSGIMLTFLMVALYCALAILRSKIMTALKEDAQVDLFDITFYFYS 180
 QY 181 LLLIQLVLSGSDSPFLFSETIHDNPPCBSSASFLSIITWITGLIVRGYRQPLESD 240
 Db 181 LLLIQLVLSGSDSPFLFSETIHDNPPCBSSASFLSIITWITGLIVRGYRQPLESD 240
 QY 241 LMSINKEDTSBOQVPLVAKMKKECAKTRKOPVKVAVSSXDPAPCKESSKYDANEVEAL 300
 Db 241 LMSINKEDTSBOQVPLVAKMKKECAKTRKOPVKVAVSSXDPAPCKESSKYDANEVEAL 300
 QY 301 IVKSPKEMNPSLFKVLVYKTFGPYFLMSFPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360
 Db 301 IVKSPKEMNPSLFKVLVYKTFGPYFLMSFPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360
 QY 361 WQGFYTYVLVFTVACLQTLVHOYFHICFVSGMRKTVIGAVRKALVITNSARKSSTV 420
 Db 361 WQGFYTYVLVFTVACLQTLVHOYFHICFVSGMRKTVIGAVRKALVITNSARKSSTV 420
 QY 421 GEIYNLMSVDQRFMDLATYINMISAPLOYILALYLMNLGPSVLGAVVMVLMVFN 480
 Db 421 GEIYNLMSVDQRFMDLATYINMISAPLOYILALYLMNLGPSVLGAVVMVLMVFN 480
 QY 481 AVMAKTKTYQVAVHMKSKDNRIKLMNEILNGIKVLKLYAMELAFKDXVLAIROBELKVLK 540
 Db 481 AVMAKTKTYQVAVHMKSKDNRIKLMNEILNGIKVLKLYAMELAFKDXVLAIROBELKVLK 540
 QY 541 KSAIYLSAVGTFTWCTPFLVALCTPAVYVITDENNILLDAQAFVSLALFNILRPFLNLP 600
 Db 541 KSAIYLSAVGTFTWCTPFLVALCTPAVYVITDENNILLDAQAFVSLALFNILRPFLNLP 600
 QY 601 MVISIYVAVSLSKRLRIFLSHEBLEPDSIERRPVKGOGGNSIVRATFTWASDPT 660
 Db 601 MVISIYVAVSLSKRLRIFLSHEBLEPDSIERRPVKGOGGNSIVRATFTWASDPT 660
 QY 661 LMGITFSIPBEGALVAVVQVQVGGKSSLLSALLAEMDKYEGHVAIKGSVAVYVPOQAMIOND 720
 Db 661 LMGITFSIPBEGALVAVVQVQVGGKSSLLSALLAEMDKYEGHVAIKGSVAVYVPOQAMIOND 720
 QY 721 SLRENILFGCQLEBYRYSVIOACMLPDLLEBSGDRTEIGEKVANSKGGOKORVSLAR 780
 Db 721 SLRENILFGCQLEBYRYSVIOACMLPDLLEBSGDRTEIGEKVANSKGGOKORVSLAR 780
 QY 781 AVYSNADYYLFPDDPLSAVDAHVKHIFENVIGPKGMLKNKRIILVTHSSYSLPOVDVILV 840
 Db 781 AVYSNADYYLFPDDPLSAVDAHVKHIFENVIGPKGMLKNKRIILVTHSSYSLPOVDVILV 840
 QY 841 MSGGKISEMGSYOELLARDGAFAEPLRTYASTBODABENGVTGVSFGPKAKOMENG 900
 Db 841 MSGGKISEMGSYOELLARDGAFAEPLRTYASTBODABENGVTGVSFGPKAKOMENG 900
 QY 901 LVTSBAGQOLROUSSSSSYSGDISRHNSITAELOKAEKKEETKMLMEADKAOGTGYKL 960
 Db 901 LVTSBAGQOLROUSSSSSYSGDISRHNSITAELOKAEKKEETKMLMEADKAOGTGYKL 960
 QY 961 SVYMDYMAKIGLFTSFLSIPLFCMCHVASLASNYMLSLWTDPTINGOEHKVALSYG 1020
 Db 961 SVYMDYMAKIGLFTSFLSIPLFCMCHVASLASNYMLSLWTDPTINGOEHKVALSYG 1020
 QY 1021 ALGISGQIAVEGYSMAVSIIGIILASRCLHVDLHLSILSPMSFPERTSGNLYNFSKEL 1080
 Db 1021 ALGISGQIAVEGYSMAVSIIGIILASRCLHVDLHLSILSPMSFPERTSGNLYNFSKEL 1080
 QY 1081 DTVDMSIMEVIMKMGSLFNVIAGACIVILATPIAIIIPGLIYFFVQRYVASSROL 1140
 Db 1081 DTVDMSIMEVIMKMGSLFNVIAGACIVILATPIAIIIPGLIYFFVQRYVASSROL 1140
 QY 1141 KRLBSVSSPYVSSHNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
 Db 1141 KRLBSVSSPYVSSHNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYPSIVANRWLA 1200

QY 1201 VRLCEVGNCTVLPALPAVIRSRHSLAGLVGSYSISLOTTTYINMLVRSSSEMETNIVA 1260
 Db 1201 VRLCEVGNCTVLPALPAVIRSRHSLAGLVGSYSISLOTTTYINMLVRSSSEMETNIVA 1260
 QY 1261 VERLKEYSETEKEAPWQIQETRPSSWPQVRVEFRNYCLRYRBDLPVLRHIVNTINGG 1320
 Db 1261 VERLKEYSETEKEAPWQIQETRPSSWPQVRVEFRNYCLRYRBDLPVLRHIVNTINGG 1320
 QY 1321 EKVGIIVGRTGAGKSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPODPVL 1380
 Db 1321 EKVGIIVGRTGAGKSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPODPVL 1380
 QY 1381 SGLRPMULDPSSQYSDBEWTSLELAHLKDFVSAIPDKLDBECAGGENLSVGORVYCL 1440
 Db 1381 SGLRPMULDPSSQYSDBEWTSLELAHLKDFVSAIPDKLDBECAGGENLSVGORVYCL 1440
 QY 1441 ABALLRKTLIVDEATPAVDLEETDDLIQSTIRQFEDCTVLTARLNTIMDYTRYVL 1500
 Db 1441 ABALLRKTLIVDEATPAVDLEETDDLIQSTIRQFEDCTVLTARLNTIMDYTRYVL 1500
 QY 1501 DKGEIOEYGAPSDILLQORGLFYSMAKDAGLV 1531
 Db 1501 DKGEIOEYGAPSDILLQORGLFYSMAKDAGLV 1531

RESULT 2
 JE0336
 canalicular multispecific organic anion transporter - human
 C:Species: Homo sapiens (man)
 C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
 C:Accession: JE0336
 R:Uchiyama, T.; Hinojosa, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Tsh, S.; Furukawa, M.;
 Bloch, B. *Biochem. Res. Commun.* 252, 103-110, 1998
 A:Title: Isolation of a novel human canalicular multispecific organic anion transporter,
 t.
 A:Reference number: JE0336; MUID:99032812; PMID:9813153
 A:Accession: JE0336
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1527 <UCH>
 A:Cross-references: UNIPROT:O15438; UNIPARC:UPI0000169859; GB:AF083552
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
 C:Keywords: ATP
 F:1306-1499/Domain: ATP-binding cassette homology <ABC2>

Query Match 45.3%; Score 4484.5; DB 2; Length 1527;
 Best Local Similarity 56.6%; Pred. No. 4.7e-236;
 Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;

QY 8 SADSDPLMDNVTWNTSNPDFTKCFONTVAVWVPCFYMACPFFYLYLSRHDRGYIQM 67
 Db 7 SGEIGSKFMDGNSLVHTENPDLTFCQNSILAWPCILVVALPCYLLYLRRHGGYIIL 66
 QY 68 TPLAKTKALGFLMIVYADLFFSFWERSGIFLAPFVLVSPILLGITTLATFLIOLE 127
 Db 67 SHLSLKKVLDVLLMVCVSMADLFYSFGLVGRAPAPVFTPLVVGWTLALATLLOYE 126
 QY 128 RRGVSSGIMLTFLMVALYCALAILRSKIMTALKEDAQVDLFDITFYFYSLLIOVL 187
 Db 127 RLGQVSSGIMLTFLMVALYCALAIFRSKILAAKGEISIPFRTTYIHVALYSLTI 186
 QY 188 LSCFSDRSPFLSETIHDNPPCBSSASFLSIITWITGLIVRGYRQPLESDLSLNKE 247
 Db 187 LACREKRPFFSAGNVDPNPPERSAGLSLFLFMFWFKMAIYGRHPLBEEKDLSLKEB 246
 QY 248 DTSEQVPLVAKMKKECAKTRKOPVKVAVSSXDPAPCKESSKYDANEVEALVKSQK 307
 Db 247 DRSGVTVQQLLEARKKQKOTARH-----KASAPGK-----NASGEDEVLLGABRP 294
 QY 308 EMNGLSFVLYKTTGYPYFLMSFPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPDMQGYRYT 367
 Db 295 R-KDSFLKALATRGSSFLISACFKLIQDLISFNPOLSLIRISNPMGDSMWGFLVA 353

368 VLFVTAQLQTLVHOFHICFVSGBRITKAVIGAVRKALVITNSAKSSTVEIYVLM 427
354 GLMFLCSMMSGLIOLHYHYHIFVGVKRTGIMGYIRKALVITNSVRASSTVEIYVLM 413
428 SVDAQRMDLATTYNMISAPLQVITLALYLMNLGSPSLAGVAVMLVPMVAVMAMKT 487
414 SVDAQRMDLAPFLNLMASAPLQIITLALYFLMNLGSPSLAGVAVMVLITPLNGAVVM 473
488 KTVAVAHMKSMDNLIKMLNELNGIKYKUYAMELAPDKYLAIRQELKYLKKSAYLSA 547
474 RAEVGVKMKLSDSKMSKSEILNGIKYKUYAMELAPDKYLAIRQELKYLKKSAYLSA 533
548 VGFVTVCTPFLVACTFAVAVTIDENNIIDQATFVSLALENLIRPILNIPWISIV 607
534 TTTTWMCSPLVLTILMAYVYVDPNVNDLAKAFVSLSIFNLRLPLNMLPDLISMT 593
608 QASVSLKRLIFLSHEELPDSIRRRPYKGGGINSITRNATFTMARSDPPTNGITFS 667
594 QASVSLKRIQOFLSQEELDPQSVERTKISPG--YAITHSCTFWAQDLPPTLHSDIQ 650
668 IPEBALVAVVQVCCGKSSLSLALAEKVEGHALKGSVAVYPOQAMIONDSLRNITL 727
651 VPKGALVAVVQVCCGKSSLSLALAEKVEGHALKGSVAVYPOQAMIONDSLRNITL 710
728 FGCOLLEPYRSVIOACALPDLILPSGDRTEIGEKVNLGGOKORVSLARAVSNAD 787
711 FGKLNKRKRYOQTLAEACALLADLEMLPGDOTEIGEKVNLGGOKORVSLARAVSNAD 770
788 IYFDDPLSAVDAAVHKHIFENYVIGPKMKLNKRLITVTHSMYLPQVDVIVVSGKIS 847
771 IFLDDELPSAVDSHVAKHIFDHVIGPEGVLAKTRVLYTHGISLPQDPIFVLADQVS 830
848 EMGYSQELIARDAFAEFLRTYASTEOBQAEKNGVTVSGPGEAKOMENGMT----- 901
831 EMGYPALTLORNGSFANFLCNVADDEQGHLEMTALKEADLEKALIBETJLSNHTDLT 890
902 ----VTDSAGKOLROSLSSSSSGD-----ISRHNHSTVE-LQKAEKKEETMKLEAD 951
891 DNDPVTYVVOQKQFROLSALSS-DGEGQKRPVPRHIGPSKAVYTEKADGA--LQOEB 947
952 KAQGVKLSVYMDYMKALIGLISFLSIFLFCMCHVSNALSNVMSLTMDPIYNGTOEH 1011
948 KALIGTELVSFMDYAKAVGLCTTALCLLVGQSAALIGAVMLAMTNDAMADSNQNN 1007
1012 TKRRLSYGALGISOGLAVFGYSMAVSGIGILASRCLHDLHLSILSPMSFERTPSGN 1071
1008 TSRLGYAALGIIQGFVYMLAAMAAAGCIGQARVHLQALHNKIRSPSFPPTPSGR 1067
1072 LVNRFSEKLDVTDSIMEPEVIMFMGSLFNVIGACIVILLAPLAIILIPGLYFFVOR 1131
1068 ILNCFSKDIYVVDVILAPVIMLNLSPFNALSTLVIMASTPLTVVILPLAVLYTLVOR 1127
1132 FYVASSRQLRLSESVSRSPVSHNETLIGSVIRAFEEORFIHQSDLKDEQKAYYP 1191
1128 FYAATSRQLRLSESVSRSPVSHNETLIGSVIRAFEEORFIHQSDLKDEQKAYYP 1187
1192 SIYANRLAVLBEVCNCIVLPAALFVAVISHSLASAGLVGLSVSYLQVTTYLVNMLVMS 1251
1188 YIISNRLSIGVEFVNCVVLPAALFVAVISHSLASAGLVGLSVSYLQVTTYLVNMLVMS 1247
1252 SEMETNIVAVERLKEYSETEKEAPMOQOETAPPSMVQVGRVFRNATCAIYRELDVLR 1311
1248 SDLESNIVAVERLKEYSETEKEAPMOQOETAPPSMVQVGRVFRNATCAIYRELDVLR 1307
1312 HINVTINGEKGIVGRTGAGKSSITLGLFINSAGEIILIDGINAKIGIHLRPFIT 1371
1308 DLSLHVHGEKVGIVGRTGAGKSSITLGLFINSAGEIILIDGINAKIGIHLRPFIT 1367
1372 IIPDDPLFSGSLRMLNDPFSQYDEDEWVLSLEAHLKDFVSLPDLKDHCBEGENLS 1431
1368 IIPDDPLFSGSLRMLNDPFSQYDEDEWVLSLEAHLKDFVSLPDLKDHCBEGENLS 1427
1432 VGGQVLCLARALLRKTILVLDATAAVDELTDDLGSTIRTFQEDCTVLTIAHRLNTI 1491

1428 VGGQVLCLARALLRKTILVLDATAAVDELTDDLGSTIRTFQEDCTVLTIAHRLNTI 1487
1492 MDYTRVVLVDKGEIQEYGAQSDLIQORGLYSNAKADGL 1530
1488 MDYTRVVLVDKGEIQEYGAQSDLIQORGLYSNAKADGL 1526
RESULT 3
571841
multidrug resistance protein, canalicular - human
C/Species: Homo sapiens (man)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
R/Accession: S71841, S71840
R/Koenig, J.; Keppler, D.
submitted to the EMBL Data Library, August 1996
A/Reference number: S71841
A/Accession: S71841
A/Molecule type: mRNA
A/Residues: 1-1545 <KOE>
A/Cross-references: UNIPARC:UPI00001746CD; EMBL:X96395; NID:g1507819; PID:CAA65259.1;
R/Buechler, M.; Koenig, J.; Brom, M.; Kattenbeck, J.; Spring, H.; Horie, T.; Keppler, D.
J. Biol. Chem. 271, 15091-15098, 1996
A/Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
A/Reference number: S71839; MUID:96279006; PMID:8662992
A/Accession: S71840
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1415-1429, 'VP', 1432-1455, 'E', 1457-1545 <BUE>
A/Cross-references: UNIPARC:UPI00001746CE; EMBL:X96395
A/Genetics:
A/Gene: GDB:ABCC2; CMOAT; ABC; MRP2; CMRP; DJS
A/Cross-references: GDB:6089489; OMIM:601107
A/Map position: 10q24-10q24
A/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C/Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
F/654-837/Domain: ATP-binding cassette homology <ABC1>
F/671-678/Region: nucleotide-binding motif A (P-loop)
F/1317-1510/Domain: ATP-binding cassette homology <ABC2>
F/1334-1341/Region: nucleotide-binding motif A (P-loop)
Query Match 36.2%; Score 3587; DB 1; Length 1545;
Best Local Similarity 47.5%; Pred. No. 5.3e-215;
Matches 746; Conservative 286; Mismatches 461; Indels 78; Gaps 20;
3 LRFGCSADGSDPLMDNVNTNTS-----NDPTKCFONTLVVWPCEYLWACPEFYFL-- 55
2 LEKFCN-----STFNSSFLDSPRADLPFCFQGYLVNHPGLFWLMLAPWQLHV 51
56 YLSRHDRGYIQMTPLNKTKTALGFLMIYCMADLFYSFWERSGIFLAPVPLVSPYTLGI 115
52 YKSRTRKSSSTTKLYLAK-QVVGFL-ILVAIEILAVLTDEDSQATVPAVRYNPSLY-L 108
116 TLLATFLDLERKRGVSSGIMLT--FWIALVCAALILRSKIMTALKEAQVDLPDIT 174
109 GTWLVLLIYSSQWQCKQNSWFLSFWILSLICGTFQFQPTLIRTLQSG-NSNLAVSCL 167
175 FYVYFSLILQVLSCGSDRPLFSETIHDNPFCSBSAFLSRITFWMITGILVNGYRQ 234
168 FFIISYGRQILILIFSAFSEN-----NSSNNPSSIASFLSITYSWDSIILKCYR 219
235 PLEGSIDMSINKEDTSEQVVPVLVKKWKECAVTRK-----QPVKVVYSSKD 281
220 PLTLDEYVMEVDEEMKTKTLVSKFETMKELQCARBALQRRQKSSQONGARLPGLNK 279
282 PAQPKSSKVDANEVEYALIVKSPKEMNPS--LFLVLYVTGPPYFLMSPFFRAIDLM 339
280 QSQSGDNLV--EDVEKKKSGTKVDVPSKMLKALFTFYMLVILKSFLLKLVNDITP 336
340 FSGPQILIKLIFKFNVDTKADWQGYFYVLLFVTAQLQTLVLAHQYHICFVSGMRKITAV 399
337 FVSPQILIKLIFKFNVDTKADWQGYFYVLLFVTAQLQTLVLAHQYHICFVSGMRKITAV 396

[illegible]

```

Db      1467 VDLEFDNLIGTITQNEFACHTYITIAHRLTITMDSKRWMLDNKXITIECSPEBELQIRG 1526
Qy      1520 LFYSNAKDAGL 1530
Db      1527 PFYFWAKKAGGI 1537

RESULT 4
571839
canaliculular multidrug resistance protein - rat
C.Species: Rattus norvegicus (Norway rat)
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C.Accession: S71839
R.Buechler, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, D.
J. Biol. Chem. 271, 15091-15098, 1996
A.Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
A.Reference number: S71839; MUID:96279006; PMID:8662992
A.Accession: S71839
A.Status: preliminary; nucleic acid sequence not shown
A.Molecule type: mRNA
A.Residues: 1-1541 <B0U>
C.Cross-references: UNIPROT:Q63120; UNIPARC:UPI000012578; EMBL:X96393; NID:G1292861; PI
C.Superfamily: human multidrug resistance protein CMOAT2; ATP-binding cassette homology
C.Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
F.100-124/Domain: transmembrane #status predicted <TM0>
F.127-151/Domain: transmembrane #status predicted <TM02>
F.160-187/Domain: transmembrane #status predicted <TM03>
F.305-329/Domain: transmembrane #status predicted <TM04>
F.354-381/Domain: transmembrane #status predicted <TM05>
F.431-451/Domain: transmembrane #status predicted <TM06>
F.456-476/Domain: transmembrane #status predicted <TM07>
F.536-564/Domain: transmembrane #status predicted <TM08>
F.574-602/Domain: transmembrane #status predicted <TM09>
F.650-833/Domain: ATP-binding cassette homology <ABC1>
F.667-673/Region: nucleotide-binding motif A (P-loop)
F.966-994/Domain: transmembrane #status predicted <TM10>
F.1018-1046/Domain: transmembrane #status predicted <TM11>
F.1104-1132/Domain: transmembrane #status predicted <TM12>
F.1203-1228/Domain: transmembrane #status predicted <TM13>
F.1313-1506/Domain: ATP-binding cassette homology <ABC2>
F.1330-1337/Region: nucleotide-binding motif A (P-loop)
F.161,1007,1010,1011/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match      35.4%; Score 3507.5; DB 1; Length 1541;
Best Local Similarity 46.2%; Pred. No. 1,4e-229;
Matches 721; Conservative 301; Mismatches 471; Indels 69; Gaps 21;

Qy      6 FCSADSDPLDMNMYWNTMNSNPDTGCFQNTVYVWVCFYIMACPPF--YFLYISRHRG 63
Db      4 FCNS---TFWMDISL--ESPEADLPICFEQTVAWIPGLFMTLLAPWOLYSVYRSRTKKS 58
Qy      64 YIQWTEINKTKYAGFLIMVVCWADLPYSFWEKSRGIFLAPVLVSPYLIGITTLIATPL 123
Db      59 SI--TFYIAKQVFVVFLILAILADISLALTEDTGQATVPVRRITNP--ILYICTLVLVIA 115
Qy      124 IQLERKQVQSSGIMLT-FWLVALVCALAIRSKIMTALKEDAQVDFRDTIFYVYFSIL 182
Db      116 VQHSRWCMCRKNKSWFLSLFWILSLVGVFOQT-LIRALLSDSKSMNAYSYLEFVSYGRQ 174
Qy      183 LIQLVLSICSDRSPLESETIHDNPPCESSASFLSRITTFWMTIGLIVNGROPLEGSDW 242
Db      175 IYLLILITAFSGP-----DSTQTPSVASPLSSITFSYVDRVLVKGKHPITLEDW 226
Qy      243 SLNK-----EDTSQVVPVLVKN-----WKKECAKTRKQVKNVYS--SKDPAQKESK 290
Db      227 DIDEGRKTRSVTSKFEAMTKDLOKARQAFORLQOKRKBATHTGLNKQSSQDVV 286
Qy      291 VDANEVEALIVKSPKKNWPSLFKVLKYTFGPYFLMSFFPKAIHDLMMFSGPQILKLI 350
Db      287 LBEAKKKSEKTTKDYPKSW---LIKSLPKTFPHVIVLMSFILKLIHDLVLFLNPLQILKLI 343
Qy      351 KPVNDTKADPMQGYFTVTLVPTACIQTLVLHQYFHICFVSGMRKIVAVIGAVYKALVI 410

```

Db 344 GFVSSNSYVWFYICALIMFAVTLIQSFCLQSYFQHCFFVLGCMCVRTTWSIYKALTL 403
Qy 411 TNSARKSTVEIYNLMSVDQRFMDLATYINMISAPLOYITLALYLMNLGSSVLAV 470
Db 404 SNLRKQYITIGETINLMSVDQKLMADATINMQLWMSYIQTITLSIFLMLRGLSIIAGV 463
Qy 471 AVMLAVPVNAVMAMKTKTYQVAHMKSKDNRIKLMNELLINGIKYKLYAMELAKDKVLA 530
Db 464 GVMVLIPVNGVLTAKIRNOVQNMKNKDKRLKIMNELLSIKILKLPAMSPSQEOVOG 523
Qy 531 IROEELVYLKKSAYLAVGFTFWCTPPLVALCTPAVYITIDENNITLDAQTAFLVSLAFN 590
Db 524 IRKRELKNTLRFQQLSLIFILQITPLVSVTFVSFVYLVDSANVLNAEKAFSITLFPN 583
Qy 591 ILRPLMLPWTSSIIQASVSVDRLEBYLGGDDLDTSIRRVSNFD---KAVKFSBAS 639
Db 584 ILRPLMLPWTSSIIQASVSVDRLEBYLGGDDLDTSIRRVSNFD---KAVKFSBAS 639
Qy 651 FTNARSDPPTLNGITFSIPREGALVAVVGVGCGKSSLSLALAEMLDVEGHVAKGSVAY 710
Db 640 FTMDPDLBATIQDVNLDKPQQLAVVGVGSGKSSLSVAMLGEMENHGHITIQGSTAY 699
Qy 711 VPQQAWIQNDLRENILFGCOLBEPYRSVIQACALPDLLEILPSGRTIEGKGVNLG 770
Db 700 VPQQSWIQNGTICKNIIFGSEYNEKKTQOVYLKACALLPDLLEILPGDMAIEGKGINLGS 759
Qy 771 GQKQSVLARAIVSNADITYLFDDPLSAVDHVGKHEBIVGEPKMLKNKTRILLVTSMS 830
Db 760 GQKQSVLARAIVSNADITYLFDDPLSAVDHVGKHEBIVGEPKMLKNKTRILLVTSMS 830
Qy 831 YLPQVDIYVMSGGKISEMGSYOEELARDAFAEFLPTAS-TQEOADAENGYTGVSGP 889
Db 820 FLPPVDIEIYVAGKTIIEKSYRDLKKGVFANMKTGFMKSGPDEEATVNN----- 872
Qy 890 GKEAKQWENGLVT-----DSAG-----KOLQROLSSSSSYGDISRHNNSTAEIQ-- 935
Db 873 DSEAEDEDDDGLIPMEBIPEDASLAMBRENSLARTLSRSSRSSRSGKSLAKSLKIN 932
Qy 936 ---KAQKKEBTWKLMDAKAQTCQVLSYVMDYMAIGLIFSLFSLFMCNFVSLAS 992
Db 933 NVLKEKKEVEGQKLIKKEFEVETKVFSTYKLQAVGWSLIFILLFGLNNVAIGS 992
Qy 993 NYMLSLMT-DDPIYNGT---QEHKRVLSYVAGLIGOGIAPGVSAVSIIGILLASRCL 1048
Db 993 NYMLSLMT-DDPIYNGT---QEHKRVLSYVAGLIGOGIAPGVSAVSIIGILLASRCL 1048
Qy 1049 HVDLHSLRSPMSFFERTPSGMLVNFPSKELDTVDSMIEPVIKFMGSLFNVIAGACIVI 1108
Db 1053 HGQLTLNITLAPRMFFPTTPTGRIVNRFSGDISVVDLPLQTLASMMWCPFGIAGTLMVI 1112
Qy 1109 LLATPIAIIIPYGLIYFVQRYVASSRQLKLESVSNPVYSHFNETLLGVSVIRAF 1168
Db 1113 CMATPVPAILIIPLSIYISVQVFYVATSRQLRDLDSVTXPSIYSHFSEVETGPIIRAF 1172
Qy 1169 EEOERFIHQDLAKDENOKAYRBSIVANRWLAIVLECVNGCIYVLPALLPAVVISHSISAG 1228
Db 1173 EHOORFLAMBEKQIDINQKCFSWITSNRMFLARLELVGNLWVFCSALLVIYAKTLTGD 1232
Qy 1229 LVGSLSVSYLQVTTYLMLVMSSEMETNIIVARLEKEYSETEKEAPMOIQTAPSPSP 1288
Db 1233 LVGSLSVSYLQVTTYLMLVMSSEMETNIIVARLEKEYSETEKEAPMOIQTAPSPSP 1288
Qy 1289 QVGRVERRATCLARYRELDPLVLRHINVTINGEKVGIIVGTGACKSSLTGLFRINESAE 1348
Db 1292 RHGRIQNNYQVRRPELDVLKGTICNIGKSEKVGAVGTGACKSSLTGLFRINESAG 1351
Qy 1349 GEIITIDINIAKIGLHRLRKITITIPDDPVLFSGSLAMNIDPSPQYDEEVMSTLEAHL 1408
Db 1352 GQIITIDIDIVASIGLHRLRELTITIPDDPILFSSSLAMNIDPFPKYSDEEVMSTLEAHL 1411
Qy 1409 KDFVSALPDKLIDHECAGGENTLVGQROLVCLARALLRKTKIIVLADATAAVLDLETDLI 1468

Db 1412 RSFVSGQLGLSETEGDNLSIGRQLICIGRAVLRKSKILLVDEATAVADLETDSL 1471
Qy 1469 OSTIRQFEDCTVLTIRHNTIMDYTRVVLVXGELIOEGAPSDLLQOGLFYSMAKDA 1528
Db 1472 QTTIRKESQCTVTTIAHRLHTIMDSKIMVLDNGKIVEGSPBELLSNNGSFYLMKEA 1531
Qy 1529 GL 1530
Db 1532 GI 1533

RESULT 5
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence, revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89447
R:anonymous, The C. elegans Sequencing Consortium.
Science 283, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_el
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: E89447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1494 <STO>
A:Cross-references: UNIPROT:Q20943; UNIPARC:UP1000017801E; GB:chr_X; PIDN:AAA83299.1, P
A:Gene: F57C12.4
A:Map position: X
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 33.2%; Score 3289; DB 2; Length 1494;
Best local similarity 47.6%; Pred. No. 9.3e-215;
Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;

Db 119 LAFPLQLERKQVSGIMLTFWLVALVCAALMILSKINTALKEDAOVDFRDTYVY 178
76 LAILTVACNKKKIIITTSQVTLVWLVLVVGIGIEFRYLSGFIYNEVALGIRATLYIA 135
Qy 179 FSLILQLVLSCFSDRSPLSETIHDNPPCESSASFLSITFWITGLIVRGYRPLEG 238
Db 136 FTSALAEFLFCRAD---VSDMYKSSSCPEYTAGSINLTQMTFLGVLVGNKSKLEN 192
Qy 239 SDIWSLUNKEDTSSQVAVLVKMKKCECAKTRKQPVKVYVSSKD--AQPKSSKVDANEE 296
Db 193 EDLMDINEIDKAENLIPSFQNLKPRIDEYH-QNIK-----KDBSALPKV----- 237
Qy 297 VEAIIYKSPKEMNPISFKLYLYTTPGYPFLMSFFPRALHDLMMFSGQIILKLIKPVNDT 356
Db 238 -----HPSFVLPDFTYKVTLLAGFFYKLCFMDLOFLAPQLLKOLIGFLEDK 284
Qy 357 KAPDMQGYFYTVLLFTACLOTVLHQYFHCIVSGMKRITAVIAGYAVKALVITNSARK 416
Db 285 NQVWVIGCSIVGIMFSSFLQSMFLHQYISMRLGNHVASVUTSAVYSALMINSERAK 344
Qy 417 SSTVGEIYNLMSVDAQRFMDLATYINMISAPLOVILALYLMNLGSPVLAVGAVVAVM 476
Db 345 GKTIQAVNLMSVDIQIDQIDMAPTIMLPMGAPQIFLSIYELMKPFGVALAGVULILA 404
Qy 477 VPVNAVMAMKTKTYQVAHMKSKDNRIKLMNELLINGIKYKLYAMELAFKDKVLAIROEL 536
Db 405 LPVNGLIAIQMRCCQTEQMKLKDERRIKMSEILNMGKVLKLYSMERSMEMVVKIREREL 464
Qy 537 KYLKSAYLSAVGFTFWCTPPLVALCTPAVYITID-ENNITLDAQTAFLVSLAFNIRFP 595
Db 465 HILKLTSTPMAALVFSWICAPFLASISFYVYVYLDPENNVLPDLTFLVALSFDILRMP 524
Qy 596 LNLIPVVISIYQASVSLKRLRIFLSHEBELPDSIRRPYKGGGNSITVRNATPTWAR 655
Db 525 LAMVAVYGEAVQCSVSNTRLKEFFAAEEMSPQT---SISHGETDSALTEVENGLFSWSS 580

Qy	656	SDEPTINGITSPJEGALVAVVGVQGGKSSLSLTLAEMDKVBESHVAKXSVAVVPQA	715
Db	581	DEDFLREISFKIQGQLVAIVGKSGSSLSLHLALCEMKNLSGVVINENIATVPQA	640
Qy	716	WIONDSLEENLIFGCOLPEPYRVSVIOACALLPDEILPSGDRTEIGEGVNLGGQXOR	775
Db	641	WIONMSLFNNLILFNKPYDLENVEDVYKNCALKEDLANIPAGDRITEIGKGINLSGGQXOR	700
Qy	776	VSLAPAVSNADIIYFDDPLSAVDAAHVAGHIFENVIGPK-GMLKNKTRILVTHSMSTYLPQ	834
Db	701	VSLARAAYQNPDIIILLDDPLSAVDSHVAKHIFENVISSTGCLASKTRVLVTHGLTYLKH	760
Qy	835	VDVLIIVMSGGKISSEMSYOELLARCGARAFPLRTASTAHOEDODAEENGVTGSGG----	890
Db	761	CDOILVIAKEGTTSELGTVOELLNNSGAFAEFLPEELIEESKTRGRVASIGDGSQVDEIL	820
Qy	891	KEAKOMENGMLVTDSAGKOLOROLSSS-----SYSGDISRH-----HNSTAELOK	936
Db	821	RDLCQVKKCIL-----KRLESHLSQESBDEKPTSRARAEYRDSRSRVLHSPSGOHE	874
Qy	937	AEA-----KKEETWKLMEADRAQTOGVKLSYWDYMAKIGLFIISLFIEMCNHV	987
Db	875	NEALLGAISEBDVPAQENTOLIEKEFTEVKGKVEFYIAYVQASISIPITLFFPLVYVSGG	934
Qy	988	SALSNVLSLMTD-PIVNGOQHTKRLSYGALGISOGIAYVGYEMANSIGGILASR	1046
Db	935	LGISLSNFLAKLSDHAKSGNFRSSDAKMLGIVAVLNGQSFVVLIASIILITGLVLRASR	994
Qy	1047	CLHVLDLHLSIRSPMSPFEPTPSGNLVRNFRSEKELTVDMSIMEVIMKMGSLFENVIGACI	1106
Db	995	ILHAGLGINIKRSPPAFFDVPDPVIGILNRIGDIAIDRTLPDVIRHMSMTIFNVVALV	105
Qy	1107	VILLATPIAAIIIPPLGIYFPQRFYVAASSQOLRLESVSRSPYVSHNETLGVSVIR	1166
Db	1055	VIMATPWAGIAFALLSVTYFIVLRIFYISTSRQLKRLESASRSPYSHFOESIQGASSIR	1114
Qy	1167	AFEEBERITHOSDLKVDENOKRAYVSIYANRPLARLCEYGCYTLFALPAVIRSRHS--	122
Db	1115	AFGVVDNFIKOSQOVDDHLIAIYPSIVANRMLARLMEVGLIYLSAAGAAYVYRDPBG	1174
Qy	1225	LSAGLVGVSYSLOVTTYLVNMLVMSSEMETNIIAVERLKEYSETEKEAPQIOETAP	1284
Db	1175	LSAGLVGVSYSVAALNIQTILNMAVYMTSELENIIVSVERIKETVYTPRIG--NSRRLAA	1233
Qy	1285	SSWPQVGEVERNNYCLARYREDIDFYLRIHINVTINGEKVGIYRTAGAKSSITGLFIRIN	1344
Db	1233	KSMPEKGEISIKNFSVRYRPGDLVLHGISAHIAISEKVGIGRTGAKSSITLALFRII	1292
Qy	1345	ESABEELIIDGININAKIGLHDLRFKITIIPQDPVYFSGSLRNKNDLPFQOYSBEEWTSLE	1400
Db	1293	BADGSGIIDGINIANLQLEOJRSCLTITVPODPVLFSGTMKNKNDLPFSYSDSOYWEAL	1352
Qy	1405	LAHLDFPSALPBDKLHDHSCAEGGENTL-----SYGBOLOVCIARALLRKTKI	1450
Db	1353	NAHLKPFYKSLDQGEHKKISBEGENDLRPRAGSKSQFQULNFVSQHQULICLARALLRKTKV	1411
Qy	1451	LVLDEATAVLDLETDLIQSTIRTOFEDCTVLTIAHRLNTIIDYRYRVILVLDKGEIOEYGA	1510
Db	1413	LVLDEAAAVAVDETSLIQKTRIREQPFKECTVLTIAHRLNTIYVDSRLLVLVDKGRAVEFDS	1472
Qy	1511	PSDLL-QORGLFYNAKDAGLV 1531	
Db	1473	PNULANPDGIFYSAKDAANV 1494	

RESULT 6
T21219
hypochemical protein F21G4.2 - *Caenorhabditis elegans*
C|Species: *Caenorhabditis elegans*
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #next_change 09-Jul-2004
C|Accession: T21219; T24002
R|Mortimore, B.
submitted to the EMBL Data Library, October 1996

A:Reference number: Z19392
A:Accession: J71219
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1573 <MTL>
A:Cross-references: UNIPROT:Q93552; UNIPARC:UP100000823F7; EMBL:Z81016; PIDN:CAB02667.1
A:Experimental source: clone F21G4
R:Kershaw, J.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z19828
A:Accession: J24002
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1573 <MT2>
A:Cross-references: UNIPARC:UP100000823F7; EMBL:Z48621; PIDN:CAA8549.1; GSPDB:GND0028;
A:Experimental source: clone R07B1
C:Genetics:
A:Gene: CESP:F21G.2
A:Map position: X
A:Introns: 21/3; 122/3; 197/3; 276/1; 654/1; 911/1; 1067/1; 1472/2; 1537/2
C:Superfamily: human multidrug resistance protein cMOA2; ATP-binding cassette homology

Query Match 32.6%; Score 3225; DB 2; Length 1573;
Best Local Similarity 42.4%; Pred. No.2,3e-210;
Matches 685; Conservative 305; Mismatches 490; Indels 134; Gaps 26;

QY 3 LRG---FCSADGSDPLMDMNTVNTSN-PDFTKCFQNTLVVWVCFYLMACFPPEFYLYS 58
DB 9 LKGDIDAF---GRRYVD-PAVMNASTVPILSGCYGHTTLVWFP----- 47
QY 59 RHDSGYIOMTFLANKTKRALGFLMTVCNADLFYSF-----WER-----SR 98
DB 48 -----TAIVFLAPILPLAOLFRRPPEIPTRRILQIKIGLACILADSL 91
QY 99 GILPAVPE-----LVSPFLIGITTLATFLTLQERRRGVOSSGIMLTFVVAL 146
DB 92 SLFVAIVAEITFGQGPVAVDVVYPTLCLANVVALIVSCRNYGIYVSGSLFSLVFT 151
QY 147 VCAIATLR---SKMTATLKEDAQVDLPEDITFYVYFSLILQLVLSGSDRSPL-FSETI 202
DB 152 ISAIPELILYMQIYNPREANNWIDYPRCIAFLIMFPCCAEYLIHCACADASBSQYKLS 211
QY 203 HDNPPCESSASFLSRITFMWITGLIVRGYROPLEGSDLMSLNKEDTSEQVVPVLVKNW 262
DB 212 AARNPSPETTSFINRITMMWFNSLCISGVKKPPELVSDLSLNEDATSNLLVPKMYNLMD 271
QY 263 KECAKTRQPKVYVSSKDPQPKKSSKVDNBEVAL-----IVKSPQKEMN 310
DB 272 KSKKFEETAAARRIGSN--ASRTNRRRTSNDTPTPLNDOSTDDYGSVPAGQSTOK--M 327
QY 311 PSLFVKLTETGPEYILMSFFPKAIIHDLIMFSGPOLIKLLIFVNDTKAPDQGYFYVTL 370
DB 328 PSIIWTLFLMKRMVDITMTFVKLLSDVLLFCNPILIKSLIFTELEKPEPMQGVVLA TM 387
QY 371 FVTACLQTLVLIHQYFHLCFVSGMRIKTAIVAGVYRKALVITNSARKSSTVGEIVNLMSVD 430
DB 388 FFSASLSSIIILSHFYFLMYRGVTRVQTCITAAVYRKTRLRLSNAARREKTVGEIVNLMAID 447
QY 431 AORFMDLATYINMTWSAPLOVILALYILMLNGSVYLAGVAVMTLMVAVNMAKTKITY 490
DB 448 VDRFOQIIPQTMQWYSNPFQIGLALFLFLQOLGVSVFGVAVMYLLPFINVITIMIRKW 507
QY 491 QVAHHKSKDNRIKLMNETILINGIKVLKYAMELAFQDKYLAIRQELTKYLLKSSAYLSAVGT 550
DB 508 QIAQMYYDEETKKNNEVLNGIKVYKLAYAMPPEHQVIEDLRBQGLGILKKAALFRTSD 567
QY 551 FTWVCTPPLVACTFAVVVTTIDENNITLDAQTAFSALFNLRPLNLIIPWISISIVAS 610
DB 568 MLNNTASPLVALVSTPATITVIDPKNVLPPELAIFSILFNQLDRSMQSAVELITQTVVV 627
QY 611 VSLKRLRIFLSHLELPDSIERRPYKQGGGNSITVRNATTW--ARSDP--PTLNGITPS 667
DB 628 VSNRLKEFLVSEELNVAIDHR--ARD--NNDVCLCKACACISWESAHQVAPPTLTNISFS 684

```
QY IPEGALVAVGVGGCKSSLLSALLAEMDKVEGHVIAIKGVAAVYVPOQAMIONDSIRENLT 727
D 668 VNRGQLVTVVGVAGKSSMLQALMGEMEKLSGSIEMGRICVYPOQPMONNTLRQIT 744
QY 728 FGCOLPEPYRSVIAQCALPDLBITPESGDRTEIGEKVNLGGCKQKVSILARAVYNSAD 787
D 745 FGKQFDEYFSRVADACALYRDLOILPLGDNTEIGEKINLSGGKAKIASILARAVYQND 804
QY 788 IYLEDPPISAVDAVGHKII FENVIGPKMLKXKTRILVTHSMYSYPOQVVIIVNSGKIS 847
D 805 IYLLDDPMSAVDAVHQSOLFSGSVIGPEQMLRNKTRITLVNLSFLEKSDLIIVNMEKIE 864
QY 848 EMGSYQELIARDGAFAEFL-----RTYASTQEQDABENGVTGVSGGKAKQKQEN 898
D 865 YSGKYDDLM--QOGAFEBQLIECEKEERERREREMASADDENSBEGIMIGDSDFEYD 923
QY 899 GMLVDSAGKOLQROLSSS--SSYSGDISRHNNSTA--ELQKAKEAKEETWKIMEA--- 950
D 924 DVM-----ASPIIDHVLGTHSHSTVSGIINRRRISTSHKORRLSTTKSHHSTISASTQ 979
QY 951 -----DKAQTGVKLSVWDYKKAIGLISPLSIFLEMCNHYVALASNTWLSLMTDDP 1003
D 980 TROLGTERVETGKMDTYKKYFGAMGMSIAVLFVLCMTTSTJFSMGRNMLTDMSDN 1039
QY 1004 IV-----NCTQHTKVRSLVYALGISOIAVFGSMAVNSIGCIASRCLAVDLHSIIRS 1059
D 1040 AARSGNLTGQPIAIRLGVVAGLGPSEIILFIQMLSLLYGVSASRNLHAPLKRNLKRV 1099
QY 1060 PMSEFERTPSGNLNVRFSEKELDTVDSMIPEVIKMFMSGLFVNVIGACIYILATPILAI 1119
D 1100 PMAYYDTPFRILINRICKOIEYDVLFPVNOFFAOCLLOVYSTLIITIMSTFVGVIV 1159
QY 1120 PPLGLIYFVQRFYVASSRQKLEESVRSFVSHFNETLLGVSVIRAFEOERFIHQSD 1179
D 1160 IPLSVMLVMVRYYIATSROKLESITRSPYSHLSIESIGSATIRAYHLVDRFCKLSE 1219
QY 1180 LKVDENOKAYYPSIVANRMLAVLECVGNCLVLPALFVVISRSLASGLVLSVSILO 1239
D 1220 TKVDSHVQCRVLANVARNMLSVRLFIQNCIVLPSALFALTRITTSVIGLVSVALN 1279
QY 1240 VTYLNLVLRMSSEMETNIVAVERLKEVSETEKEAPMOIE-TPAPSPWQGVVERPNY 1298
D 1280 ITTYLANAVROITLERNIVSEVKEYAETETEMKSEPKERPQWMBEGKIWMANY 1339
QY 1299 CLARYEDLDVLRHINVTINGSERVGIAGRTGAKSSLTLLGFRINSAEGEIIIDGINI 1358
D 1340 SARVRPGINLVVQOLNVEIKPHEKGIYGRGAGKSVTSLFRILIEAEGQIVVDGINTL 1399
QY 1359 AKIGLHLRPFKITIIPQDPVLFSSSLRNLDLPFSGQYSPBEVWTSLELAHLKDPYALBDK 1418
D 1400 AEIGLHLRSNLTIIPODPVLFSGTLRFNLDPFNHYSQDITWKILEMANLEFATANEQ 1459
QY 1419 LDHECAEGENLSVGOQLVCLARALRKTIILVLEDEATAVDDETDLIOSTRTOFED 1478
D 1460 LMYITIGSGNISVGOQLVCLARALKRTRVILLDERTAAVDVSTDALIOKTRIEEPAN 1519
QY 1479 CTYVLIARLNTIMDYTRVIVLDKGEIOEYGAPEDDLQORGL-FYSNAKDAGLV 1531
D 1520 ATVTLIARLNTIMDYDRIIVLDGKVGGEFDSPANLLSNRSEFYNAKRAGLI 1573
RESULT 7
T42216
multidrug resistance-associated protein homolog MLP-1 - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 03-Dec-1999 #sequence_reviseion 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T42216
R:Hirohashi, T.; Suzuki, H.; Ito, K.; Ogawa, K.; Kume, K.; Shimizu, T.; Sugiyama, Y.
Mol. Pharmacol. 53, 1068-1075, 1998
A:Title: Hepatic expression of multidrug resistance-associated protein-like proteins ma
A:Reference number: Z22081; MUID:98279126; PMID:9614210
A/Accession: T42216
```

```
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1502 <HR>
A:Cross-references: UNIPROT:O88269, UNIPARC:UPI000012575A, EMBL:AB010466, NID:93242457,
A:Experimental source: strain Sprague-Dawley, liver
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
Query Match 32.3%; Score 3195.5; DB 2; Length 1502;
Best Local Similarity 43.1%; Pred. No. 2,2e-208;
Matches 655; Conservative 290; Mismatches 525; Indels 51; Gaps 9;
QY 22 WNTSNPD-----FTKCFQVTVLVWVPCFYLMACFPYFYLYLSHDDGYIOMTPLYNKT 75
D 20 MNQTEQRPVAYHLLNLCFLAAGSVVPMYIMVGLPIYLYLHHNGCCYLRMSLFFKIM 79
QY 76 ALGFLMIVCMADLFYSFWERSRGIPLAPVFLVSPYLLGTTLLATPLQLERRKQVSS 135
D 80 VLGFALILLVFNAVPLMRHRCMPQAPPELLIPVWMLTMSFATPLIHMERKGVAS 139
QY 136 GIMLTPVLVLCALILRSKIMTALKEDAQVDFRITPFYVYSLILQVLVSCFSDRS 195
D 140 GLRFGYV---LLCCLVPAIDTVQOASGFRQEBLHILATYLCISLVVAELVLSCLVDQ 196
QY 196 PLFSETIHDPPCPESASFLSRITFMWITGLIVRGYROPLESGDLSLNKEDTSBOVVP 255
D 197 PFSBEDKPLNCPCEAASFPKAMFWAGSLKIGTRKLGPRCDLSLERNSEELVS 256
QY 256 VLVNWKKECAKTRKQPVKVVYSSKDBAOPKESKVDANEVEALIVKSPQKWNPSLKF 315
D 257 QLEBERNRN-----PSELRGHGHGSM--GTRETAFL--QEPERSQGRPLLR 299
QY 316 VLKTPGPYFLMSFFFAIHDLMMFSGPOLKLIKLVNDTKADWQGYFTYLLFYTAC 375
D 300 AIKWFVSTPLGLTLVVIDARFAVPRKLSLFLFEMGDLSESAWGMVLAVLMFLSAC 359
QY 376 LQTLVHQYHICVSGMRIKTAVIGAVYKALVITNARSKSTVGEIIVNLSVADARFM 435
D 360 LQTLFEOQYMRVKVLOMRIRKTAITGLVYKVLVLSGSRSSKSAAGVNVLSVDVRLV 419
QY 436 DLATYINMISAPLOVIALYLLMLNLPSPVLGVAVMVLPVNAVAMKTKTYQVAHM 495
D 420 ESILHNLGWLPLMIIVCVVLMQLLGPRBALTVANFSLPLNPFITTKRSHQBEOM 479
QY 496 KSKDNRIKAMEIINGIKVLKLYAMELAFQKVLAIPOBELKVLKKSAYISAVGTFTWC 555
D 480 ROKASRBRLTSSMLRTVRIKSHGWECAFLERLHITIGOEIGALKTSAFLFSVLSVFOV 539
QY 556 TPFLVALCTPVAVYVTTDENNIILDAQTAFFVSLALFNILRPLNLTLPWYISSIVQASVLR 615
D 540 STELVAVPRAVHVLVAEDNAMDEKAFVTLVLSLNKQALFSPSHCICLQARVSPDR 599
QY 616 LRIFLSHEBELPDSIERRPVKGQGTNSITVRNATPTWASDEPTLLNGITFSIPGALVA 675
D 600 LAAPFLCLEEVDPMGMVLSPPR-CSSKORISIHNGTAMQGESPCGHGIMLVPOGCLLA 658
QY 676 VVGQVGGKSSLLSALLAEMDKVEGHVIAIKGVAAVYVPOQAMIONDSIRENLTREGQLEP 735
D 659 VVGPVGAKSSLLSALLGELLKVEGVSIEGSAVYVQBAVMVQNTSVENAVCFQOEIDL 718
QY 736 YNSVIOACALPDLBITPESGDRTEIGEKVNLGGCKQKVSILARAVYNSADYILPDP 795
D 719 WLQEVLEACALGSDVASFPAQVHTPVGEQGMNLSGGCKQLSLARAVYRRAVYLLMDPL 778
QY 796 SAVDAVGHKII FENVIGPKMLKXKTRILVTHSMYSYPOQVVIIVNSGKISEGYSQEL 855
D 779 AALDAVHSGVFRKQVIGSGLQGTTRIIIVTHLHVLPQADQILIVLANGTIAEGSYQDL 836
QY 856 LAADGAFAELRYIYSTEQDABENGVY-----GVSGGKAKQKQENMLVTDSAGKQ 909
D 839 LHRNGALVGLLDGAROPAGGEEBAHAATSDLDGGSGGCTPTPRRE----- 886
QY 910 LQROLSSSSYSDISRHNNSTAELOKAKEAKEETWKLMEADRAQTCQVLTIVWDYKMA 969
```


Db 1088 DSIRSPPVYAOFGELNGLSTIRAYKAYDRMADINGRSMNNIRFTLVNMGANRMLGIRL 1147
Qy 1204 ECVGNICIVLFAALFAVI-----SRHSLSAGLVGLSVSYSLQVTTYLNMVMSSEMETNI 1258
Db 1148 ETLGLMIMLWLFASFAVMONGBAENQAFASWTGULLSTVALNITSLTGVLRLASLAENSL 1207
Qy 1259 VAVERLKEVSETKEAPMOIQETAPSSMPVOGRVEFRNYCLIRREDLDFVLRHINVTIN 1318
Db 1208 NAVEKVGVIYIIPREAPFVIEENRPPCMPSGSIKFDVVRKYRPOLPVYLHVSFPIH 1267
Qy 1319 GGEKVIIGRTGAGSSLTGLFRINESAGEIIIDGINIAKIGHADRKFRTIIPOPV 1378
Db 1268 PTDKVIIGRTGAGSSLLNLFRIVEYKGRILLIDCCVCGKFGMLDKRGILGIIPOSPV 1327
Qy 1379 LFGSLRKMLDPPFSQYSDIEWTSLLEALXDFVSALPDKLDHECAGEENLSVGORQV 1438
Db 1328 LFGSVIRNLDPPFGHNNADLWESLERHAKOTIRRNPLGDAVSEAGENSVGQRLL 1387
Qy 1439 CLARALKRKILVLDEATAVDETDLIQSTIRTOEDCTVLTIANRLTIMDYTVI 1498
Db 1388 SLSRALLRSKILVLDEATAVAVDTDALIQTIREEFKSCMTLIIAHLMTIIDCKIL 1447
Qy 1499 VLDKGEIOEXGAPSDLLQGRGLFVS-MAKDAG 1529
Db 1448 VLDSGRVQFSSPEMLLSNEGSSFSKWOQSTG 1479

RESULT 12

D86428
glutathione S-conjugate transporting ATPase (AtMTP1) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: D86428
R/Theologian: A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasey, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hulcar, L.
Nucleotide 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Xu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86428
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1622 <STO>
A:Cross-references: UNIPROT:Q9C8G9; UNIPARC:UPI000009CF47; GB:AE005172; NID:G11055814; F
C/Genetics:
A:Map position: 1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 23.8%; Score 2356.5; DB 2; Length 1622;
Best Local Similarity 35.6%; Pred. No. 2.9e-151;
Matches 569; Conservative 274; Mismatches 564; Indels 193; Gaps 31;

Qy 47 WACPEPF-YFLVLSRDRGIVOMTPIKTKTALG-----FLIMTVICADLFYSPW 94
Db 8 WYCKPEVNGVMTKVTVYAGATVPCALDSFVLGISHLVLLILCLTKRLMLITYTDHKVYKFC 67
Qy 95 ERSR--GIFLA-----PVFLVPTLLGTTT-----LATFLIQERERKQVS 134
Db 68 LRSKMFSLFLLALAYATAPLFRLL--VMRISVLDLDGAGPPPEAFMLVLE--AFAMG 122
Qy 135 SGIMLTF-----WLVAVCALALIRSKIMTKALKEDAQVDLFRDITVYVFSLL 183
Db 123 SALVMTVETKTYIHLEIRMYVFAVIYALVGMVLLNLVLSYK--EYVGSFKLYLISEVA 181
Qy 184 IQLV-----LSCSDRSPLFSETHD-----PNPCESASPLSLITPMW 223
Db 182 VQVAFGTLTFVYFPMLDPPYGYTVPGTENSEDEYEEELPGGENICPERHANLFDISIFSW 241

Qy 224 ITGLIVGYPRLGSDLSLNKEDTSEQVPPVLYRNMKKCAKTRQPVKVVYSSKDA 283
Db 242 LNPMTIGSKRPPLTEKQVWMLDITDKTETLRSFOKSMDEK----- 285
Qy 284 QPKESKVDANEVEALIVKSPKEMNPSPLEKVIYKTFGPFLLSPFFKALHDMFMFGP 343
Db 286 -PK-----PMLRALNLSIGKFFMGMGFFKIGNDGSGQFVGP 320
Qy 344 QILKLLIKFVNDTPAPMGOGYFYTVLLFVTACIQTVLHQYFHI CFVSGMR IKTAVIGAV 403
Db 321 LLNLELLKSM-QLNBPAMIGIYVAISIFVGVLVGLCAQGFQVNMRYGRLRSALIAAV 379
Qy 404 YRKALVITNSARKSTGEIYNLSVDAQRMDLATIYNMTSAPLOYIALLYMLNLG 463
Db 380 FRKSLRITNEGRKRFQTKITNMTTAAESLQOQCSLHTMSAPFRITVALVLLYQDLG 439
Qy 464 PSYLAGVAVWVLAVPVNAVMMKTKTKYQVAMKSKDNRIKMLNELINGIKYLKYAMELA 523
Db 440 VASIIIGALFLVLMPIQTVIISKQKLTKEGLQRTDKRIGLMEVLAAMDVVKCYAMENS 499
Qy 524 FKDKVLAIROBELKVKLKKSAVLSAVGTFTWCTPEFLVALCTPAYVYTI DENNII DAOTAF 583
Db 500 FQSKVQVTRDELSPFRAQLLSAFNMFILNSIVLTVTVSGVFSLLGCD--LTPARAF 557
Qy 584 VSLAFNLRPLPLIIPMVISIYQASVSLKRLAIFLSHER--LEPDSIERPKVDGGT 641
Db 558 TSLSLFVLRPPLPMLPVIITQMVNANVSLNREEVSLTEERVLIPND---PIEP--GQ 611
Qy 642 NSITVRNATFTW-ARSDPTLNGITPSPREGALVAVVQVCGKSSLSALLAEMD-KVE 699
Db 612 PAISTIRNGYFMSDRADEPFLSNINLDPGLSLAVAVSGTSGKTSLSANLGLIPARSD 671
Qy 700 GHVAIKGSVAVYPOAMVIONDSURENLIIFGCOLEPPYRSYIOACALPDELIPSGDRT 759
Db 672 ATVTLRGVAVYPOVSWIFNAVTARNLIFGAPRPOEKERYIDVTALQHDLELLPGDGLT 731
Qy 760 EIEGKVNLSCGQQRVSLAAVYNSADIFYFDPELSAVNDVHGKHFENVYIPGCMYKN 819
Db 732 EIEGKVNISGQQRVSMARAVYNSVNCIIDLPLSLADHVAQVVEKECI--KRELGQ 789
Qy 820 KTRILVTHSMYLVQVNDVIYIMSGSKISEMSYOELLARDGAPAEFLRYASTEEOQDAE 879
Db 790 TTRVLTYNQLHPLSQVDKILVHEGTYVEBGTIEBLCHSGPLFORLNMAGKVEDE--SE 847
Qy 880 ENGVTVSGPKEAKOMENGMVYDSAGKQRLQRLSSSSSYSGDISRHNNSTAELOK--- 936
Db 848 ENGEAEVD--QTSYKVPVNG-----NANNLQKGI 875
Qy 937 -AEAKKEETWLMEDAKQTCQVYKLSYWDYMKAI-GLFISFLSIFLMCHVSLASNY 994
Db 876 ETKNSKEGNSVLYRKEERETGVSMKVLERYQNALGAMVMMVLVTCVLTQVFEVSSST 935
Qy 995 WLSLMTDQ--PIVNGTOHTKVRLSVYGALISQGINAFVSGMAVSIIGIIASRCLAHDL 1052
Db 936 WLSEWTDGCTKTHGPLEFYN---IVYALLSFGQVSVTLINSYWLIMSKSLTAACKWHDAM 991
Qy 1053 LHSILRSPMSFFERTPSGNLVNRFSEKELDYDSMIPVIYKAFMGSIFNVIGACTIYLAT 1112
Db 992 LGSITRAWVFQNPPLGRITINRPAKMGDDIDRYAYAVNMFMGSIADLGLTILIGIVS 1051
Qy 1113 PIALIIPPLGLIYFFQRFVYASRQLKRLSEVSRSPYSHPNETLLGVSIVIRAEBOE 1172
Db 1112 RMAELINGSMNNNIRFTLVNMAANMIGIRLEVLGGLVWMLTASLAVWONGKAQAQOAYA 1171
Qy 1228 GLVGLSVSYSLQVTTYLNMVMSSEMETNIIVAVRLKEVSETKEAPMOIQETAPSSM 1287
Db 1172 STMGLLSYALSTISSLTAVALRLASLAENSVARVGNVYIPIBEADVLVIEENRPPGM 1231
Qy 1288 POGVGRVEFRNYCLIRREDLDFVLRHINVTINGEKVGIVRTGAGKSSLTGLFRINESA 1347

Db 1232 PSSGSIKEDVLRKRPFLVHLGVSLIPMDKVGIVGRTGAGKSLNALPRIVLE 1291
 Qy 1348 EGEIITIDGINAKIGHDLRFRKTIIPDDPVLFGSSLMNIDPFSQSDDEWMTSLIAH 1407
 Db 1292 KGRILIDECIDIGRGLMDLRKVLGIIPQAPVLFSGTVAFNIDPSEHNDALMWSLEBAH 1351
 Qy 1408 LKDEVSALPDLDHBCAEGENLSVGRQLVCLARALRKTIIVLDEATAVLETDL 1467
 Db 1352 LKDTIRRNPLGLDEVTAGENFVSQROLSLARALRKRSLILVDEATAVDRIVL 1411
 Qy 1468 IQSTIRTOFEDCTVLIJAHRLNTIMDYTRVIVLKDGEIOEYCAPSDILQORGLFYSMAKD 1527
 Db 1412 IQKTRBEFKSCVMIILAHRLNTIIDCDKVLVDSGKQEFSSPENLISNBSSEFSK--- 1468
 Qy 1528 AGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAEPRTRE 1567
 Db 1469 -----MVOSTGTANAEYLRSLTLENK-----RTRE 1493

RESULT 13

E86428
 Probable ABC transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: E86428
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizlar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, Y.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroso, J.S.; Maiti, R.; Matzalli,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86428
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-1495 <STO>
 A:Cross-references: UNIPROT:Q9C8H0; UNIPARC:UPI00000482FF; GB:AE005172; NID:g11055818; F
 C:Genetics:
 A:Map position: 1
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 23.5%, Score 2326; DB 2; Length 1495;
 Best Local Similarity 34.3%, Pred. No. 3e-149;
 Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;

Qy 2 ALRGFCG-ADGSDPLMDNVTNNTSNPDFTKCFONTVLVWP-----CFY-LMACFP 51
 Db 5 ALMNYCPVADGF-----MEKAVDGAFGAYTPCAIDSLVMLVSHFVLGLGCFYRIWIIF- 58
 Qy 52 FYFYLISRHRGVIQMTPIKTKALGFLMIVCMAD-----LFYSFWEWSRGIPLAPV 105
 Db 59 -----HNTAQIYVRKKYKVCVLG-LIACYCVEEVLRLVNGISFLPMDDEBTPPPF 110
 Qy 106 FLVSPILLGITTLLATEFLIQLERRKGVSSGIMLTFLMLVALCALAIRSKIMALKEDA 165
 Db 111 EVASLWAEAFWFMVLIIGLETQYKQYKPRWYVRFGLVYLAVADAVLDLVLKXIN 170
 Qy 166 QVLDLFRDIT--FVYVSLLIQLVLCFSRSPFLFSETIHP-----NCP 209
 Db 171 RTALYLISSRCQALFGILLIYIPEL--DPYGYHVNNEPILDNVEYDALRGEBHICP 228
 Qy 210 ESSAFSFRITFMVITGLIVGKQPLSGDMLSLNKEDTSEGVVPLVKKMKKECAKTR 269
 Db 229 ERASISIRIYFGITPLMQIGYKPKITEKDWQIDKMDQETELIKRFORMTESRR-- 286
 Qy 270 KQPKVYVSSKDPQPKSSKVDANEEVALIVSPKKNPSSLFKVLYKTFGYPFLMSF 329
 Db 287 -----PK-----PMLRLRLNNSLIGRFMLAG 307

Qy 330 FFKAHLMWSSGPQIILKIKFVNDTKAPDMOGYFTVILLFPTACIQTLVHOYFHICF 389
 Db 308 IFKIGNDLSQFVGVIILSHLRSMQED-PAMVGYVAFTIFVGTGLVLCDAQYFQNW 366
 Qy 390 VSGRIRTAIVAGVYRRALVITNSARKSTVGEIVINLSVDAQRFMDLATYINNIWAPL 449
 Db 367 RVGRRLSTVLAALFPHKSLRTHARKNFPASGKTNMTITDANLQOISQDLHGMSAPF 426
 Qy 450 QVIALYILMLNGLSPVLAGVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 509
 Db 427 RIIVSMILLYQQLGVASLFGSLIFLLIPIQTLIIKSKRKLTGEGIQWTDKRGVITNEIL 486
 Qy 510 NGIVALKYAMELAFKXKVALIROBEKVLKXKAYLSAVGFTVWCIPFLVACTFPAAVY 569
 Db 487 GSDTVKCYAMEKSFESRIQIRNEELSWERKQDLAFNSFILNSIPVAVTVVSFGFV 546
 Qy 570 TIDENNLIDQAFVSLALFNILRFPNILEPMVSIYQASVSLKRL-RIFLSHEBLEPD 628
 Db 547 ILGSD--LTPARATSLFLAVLRFPLMLPMLISQVYVAVVAVVAVVAVVAVVAVVAVV 604
 Qy 629 SIERRPVKDGGINSTIVRNATFTW-ARSDPTLNGITFSIPGALVAVVAVVAVVAVVAVV 687
 Db 605 NPLQF-----GTALISIKNGYFSDGKTKTPTSLDINLEIPVGTIVAVVAVVAVVAVV 659
 Qy 688 LSLALAEADKYE-GVVAIKGSVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 746
 Db 660 ISAMIGELSHAETTSVIRSVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 719
 Qy 747 LPDEILIPSGDRTEIGKGVNLSGGKQORVSLAAVAVVAVVAVVAVVAVVAVVAVVAVVAVV 806
 Db 720 QHIDLPLRGRDLTEIGRGNVSIQKQKORVSAVAVVAVVAVVAVVAVVAVVAVVAVVAVV 779
 Qy 807 FENVIGKGMKNKRIIVLTHSMYLPQVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 866
 Db 780 FDSGM--KDELRGTRVLVNVQHLFLMLKXILVSGMGIKEBTFLVLSGILFKLM 837
 Qy 867 RTASTQEDQDAEENGVTVGSGPEKAKQENGLVTVDSAGKQLQORLSSSSSYSGDISR 926
 Db 838 ENAKMDATQEVNTN-----DENILKLGPTVTVDS--ERNIGSTRK----- 876
 Qy 927 HHNSTAELOAABKEETWLMADKQOTGVLSVYWDYWKAI-GHIFISLSTIFLWCN 985
 Db 877 -----QKGRRS-VLIKQERETGILSMNVRKYEAAGSLAVVAVVAVVAVVAVVAVV 924
 Qy 986 HVSALASNYLSLMTDDPIVNGTOEHK-----VRLSVGALGISQGIAGVGSMAVSIG 1040
 Db 925 EVLRVSSSTWLSITWD-----QSTSKNYSFGYIVVAVVAVVAVVAVVAVVAVVAVVAVV 977
 Qy 1041 GILASRLAYDLSILRSFMSFFERTPSGNLVNRSFKELDTVDSMTPEVIKMFSGLEFN 1100
 Db 978 SLHAARLHDAMLSSIRAPMLFPHTPTRVIVNRFKSDIGDDIRNVANLMMFMNDLQ 1037
 Qy 1101 VIGACIVILATPIAIIIPGLIYFVQRFVYASSROKRLSESRSRPSYSHNETIL 1160
 Db 1038 ELSTPALIGVSTISLWALPMLLFLFYAALVYQSTIREVRRLDSVTRSPILYAOFGALN 1097
 Qy 1161 GVSIVTAFEEOEERFIHOSDLKVDENQKAYPSIYANMLVRLCEVCNCLVFAALPAVI 1220
 Db 1098 GLSSIRAYKAYVDRNAKINGSMONNIRFTLANSSNMWLTIRLETLOGWIMWLTRAVL 1157
 Qy 1221 SRHLSL-----AGLVGLSVSYSLQVTTYLNLVMSSEMETNIVAVRLKEYSETEKAP 1275
 Db 1158 QNGTNNQAGFASMTGMLTSTYTNITSLSGVLRQASRAKNSINSVERVGVYIDLPSBAT 1217
 Qy 1276 MOIQTAPSPSPVGVGEVFRANCYARRELDLFLRLINVTINGGEKVGIVGRTGAGKSS 1335
 Db 1218 DIIENRFPVCGWFGSSGSIKEDVLRKRPFLVHLGVSLIPMDKVGIVGRTGAGKSS 1277
 Qy 1336 LTGLFRINESABGEIITIDGINAKIGHDLRFRKTIIPDDPVLFGSSLMNIDPFSQSD 1395
 Db 1278 MLNALPRIVEKGRIMIDCDVAKFGLTVRRVRLSIIPQSPVLFSGTVAFNIDPSEHN 1337

Qy 1396 DEEWTSLSLAHKDPFVSALDPKLDHECAGEGENISVQOROLVCARALLKRTKLVIDE 1455
Db 1338 DAGMEALHRAHAIKDIVISRNPFGLDAEVEGEGENISVQOROLSLARALLRRSKLVIDE 1397
Qy 1456 ATAAVDLETDDLIOSTIRTOPEDECVLTIARHLNTIMYTRIVYLDKEIOEGAPSDLL 1515
Db 1398 ATASVDVDTBLLIQRTIEERKSKCTMLVIAHRLNTIICDRLVLSQGVLEIYDSPOELL 1457
Qy 1516 -QORGLPFYSMAKDAG 1529
Db 1458 SRDTSAPFRMVHSTG 1472

RESULT 14
T27408
hypotheical protein Y75B8A.26 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T27408
R:Barlow, K.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z20361
A:Accession: T27408
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1144 <WTL>
A:Cross-references: UNIPROT:O9XM61; UNIPARC:UPI000007AA2A; EMBL:AL033514; NID:e1343251;
C:Experimental source: clone Y75B8A
C:Genetics:
A:Gene: CESP:Y75B8A.26
A:Introns: 63/3; 120/3; 222/3; 271/3; 372/3; 451/1; 636/1; 687/3; 778/3; 953/2; 1042/2;
Query Match 23.3%; Score 2312; Db 2; Length 1144;
Best Local Similarity 42.5%; Pred. No. 1.7e-148;
Matches 491; Conservative 228; Mismatches 381; Indels 54; Gaps 16;

Qy 396 KTAIVAGAYRKALVITNSARKSSYIGEIVNLSVDAQRPMIATYINMISAPLOVIAL 455
Db 9 QTVSNATLHKILRLSPSARSNRGTAGEIILNHAADVIEIIVHSVPYIQMWSVFPQVTLAM 68
Qy 456 YILMNLGSPVLAVAVWVAVVAVNAVMKTKTVQVAMKSKDKRILKMLNGLIGIVL 515
Db 69 TMLATTLGMAWAGVCMILFLPLNLCSTRFKLSQOKOMKIKDKRITKLSNMLNGIKV 128
Qy 516 KLYAMELAFKDVLAIRQBELKVLKKSAYLSAVGFTWVCTPEFLVALCTFAVYV--TTDE 573
Db 129 KLYAMESEFEDQINLRKAEVAMLANVCLISRIIVANASPFVVAIGSFCTYVLMSPDE 188
Qy 574 NNILDAQTAFFVSALFNILRPLNLIPMYISSIVQASVSLKRLRIFLSHEELPDSIERR 633
Db 189 NG-LTPSVAFVALRTFNQLRQPMRWAMILNTLVQARVSNKRLRQFLNDEEME---RKT 243
Qy 634 PVKOGCGNSTIVRATRTW-ARSDPPTLNGITPISPBALAVAVVQVCGGSSILSL 692
Db 244 EVALG--NAIVFKNASLNMKGPQNPVVKDKLSATIKQGLAIAGSVGGKSSLSLAVL 300
Qy 693 AEMDEVGHVAIKGSVAAYVPOQAWIQNDISRENIIFGCLPEPYRSVIAQALPDEI 752
Db 301 DEMVLDRKVKVGGSIAYVPGHSMIFNKTIKENILFGBELSYTFDQVVGSQLKTRRH 360
Qy 753 LPSGDRTEIGEKVNLGGQKORVSLAPAVSNADIVYFDPLSAVDAAVGHIFENYIG 812
Db 361 FQGGENTWVGENGITLISGGQKARISLARAVYQDKIYLLDDPLSAVDAAVGRALPDKYVG 420
Qy 813 PKGMKNTRLIYTHSMYVLPOVDVLIYMSGKISEMGSYQGLLARDGAFALFTAYST 872
Db 421 PDGLIRSKTRVLVYTNLYOTTKVDITTYLEDOQIVQHSFEDIAVVDGFGRLMSECE 479
Qy 873 EOEOBAENGVTVGSGPKAEKOMENGMVTDGAKOLOROLSSSSSSYGDISRHNSYA 932
Db 480 SDEVDVADEBAESSEASVTPPVVLENG---DNGALIKSSQIDRINSHSEKSR----- 529
Qy 933 ELQKAERKKEETWKLMEADKQOTGOVKLSVYWDYMKAIGLFISFLIFLMCNHVSALAS 992

Db 530 ---KSEEPQKVERKVE--NVQLGRVKKSVQLYIKTMGTGINSASALFFLAHTTWIMR 584
Qy 993 NYMLSLWTD-----PIVNGTQE-----HTKVRLSYVGAIGISOGIAVFGSM 1035
Db 585 SLMLSDMSNENMAIKKATLSSVDYLNSTSSVDGVSVEYTRILIVYAGFG---GLEMLLAL 641
Qy 1036 A---VSTIGIILASCLHVDLHSLTIRSPMSFEERTPGSGNLVNRPSKEILDVDSMIPVIK 1092
Db 642 APTVLITGSLRASYGLSHSPLIHALLVAPISFFDPTPTGRIRINRLSRDLVIDK-LQDNIR 700
Qy 1093 MPMGSLFNIVIGACVILLATPIAIIIPILGLIYFPQRFVVASROIKRESVRSRPY 1152
Db 701 MCTQTLNACMILVLISTSTPIFLVCAAPLILYFPVVIYIPIYSRQURLESANRSPIL 760
Qy 1153 SHFNETLLGVAVIRAFEREOERFIHQSDLKVDENOKAVYPSIVANRLAVLECVNCIVL 1212
Db 761 STIAESTHGASIRAFDTERTTALSTNVDKFKAQCRILSISMNRMLATRIELLGNTCVL 820
Qy 1213 FAALFAVISR--HSLSAGVLGSYSLOVTTYINMLVRMSSEMETNIVAVERLKEYSET 1270
Db 821 FASLSATLSTKYFGLTQMAGLSVSVALTITREVNI CVRSSEISNIVSVERVEXOKL 880
Qy 1271 EKEAPWQOGETAP-PSMPQVGRVFRNYCLRYREDDLVLRHINVTINGEKYIVRT 1329
Db 881 EPEAPWRLEKSLNEBEKMPVKGLTELDGFSMRYNKLPVLKNDLKIEGSRIGVIGRT 940
Qy 1330 GAGKSSLTGLGFRINESAGEIIIDGINIAKIGLHDLRFKTIITIPDVLFPSSGLRNMID 1389
Db 941 GSGKSSLTLMALYRMIEGSSGITKIDVVDITIGLHQLRKLIIPQEPVVSGLTRFPLD 1000
Qy 1390 PPSQYDEEWTSLSLAHKDPFVSALDPKLDHECAGEGENISVQOROLVCARALLKRTK 1449
Db 1001 PFNQSDQIQWNCLEICQLKQFAQEDBDTLIRYIAEGKKNMSVBRQLLCLCRALLBRAR 1060
Qy 1450 ILVUDEATAVDLETDDLIOSTIRTOPEDECVLTIARHLNTIMYTRIVYLDKEIOEYXG 1509
Db 1061 IVILDEATASVDVTVGQVQARIRQHPQOSTTISIAHRLDITVDSDRIVLDAGVAEFD 1120
Qy 1510 APSDLLQORGLFYS 1523
Db 1121 TPSNLLNLPDSLYS 1134

RESULT 15
S64757
probable membrane protein YLL015W - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypotheical protein A255; hypotheical protein L1313
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S64757; S64763; PS0041; S69391; S70560
R:Miogsa, T.; Zimmermann, F. K.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64743
A:Accession: S64757
A:Molecule type: DNA
A:Residues: 1-1559 <MIO>
A:Cross-references: UNIPROT:P14772; UNIPARC:UPI000126A9C; EMBL:Z73120; NID:g1360184; PI
A:Experimental source: strain S288C
R:Goiteau, A.; Purnelle, B.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64761
A:Accession: S64763
A:Molecule type: DNA
A:Residues: 1-1559 <GOP>
A:Cross-references: UNIPARC:UPI0000126A9C; EMBL:Z73120; NID:g1360184; PIDN:CAA97460.1; P
A:Experimental source: strain S288C
R:Boy-Marcotte, E.; Damak, F.; Camonis, J.; Garreau, H.; Jacquet, M.
Gene 77, 21-30, 1989
A:Title: The C-terminal part of a gene partially homologous to CDC25 gene suppresses the
A:Reference number: PS0041; MUID:89306677; PMID:2545538
A:Accession: PS0041
A:Molecule type: DNA

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:56:46 / Search time 243.004 Seconds

(without alignments)
5594.767 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903
Sequence: 1 MALRGFCSADCSPLMDMNV.....RSVAVAKAKPKFSTSPDLSL 1927

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	79.4	1531	MRP1_HUMAN	P33527 homo sapien
2	7769	78.5	1515	09UQ9_HUMAN	Q9UQ99 homo sapien
3	7695	77.7	1531	0864R9_MACFA	Q864R9 macaca fasc
4	7689	77.6	1531	086450_MACFA	Q86450 macaca fasc
5	7444	75.2	1459	09UQ97_HUMAN	Q9UQ97 homo sapien
6	7419.5	74.9	1456	09UQ90_HUMAN	Q9UQ90 homo sapien
7	7272	73.4	1531	06UR05_CANFA	Q6UR05 canis famli
8	7242	73.1	1439	059G19_HUMAN	Q59G19 homo sapien
9	7230.5	73.0	1530	08HX05_BOVIN	Q8HX05 bos taurus
10	7104.5	71.7	1400	09UQ98_HUMAN	Q9UQ98 homo sapien
11	7002.5	70.7	1528	MRP1_MOUSE	Q35379 mus musculu
12	6932.5	70.0	1532	0810B4_RAT	Q810B4 rattus norv
13	6932.5	70.0	1532	08CG09_RAT	Q8CG09 rattus norv
14	6892	69.6	1523	0810G9_RAT	Q810G9 rattus norv
15	6301	63.6	1525	05F364_CHICK	Q5F364 gallus gall
16	6075	61.3	1215	068CP7_HUMAN	Q68CP7 homo sapien
17	4526	45.7	1522	05SUP4_MOUSE	Q5SUP4 mus musculu
18	4522	45.7	1519	080XK8_MOUSE	Q80XK8 mus musculu
19	4522	45.7	1522	056PH0_MOUSE	Q56PH0 mus musculu
20	4508.5	45.5	1523	059DU0_MOUSE	Q59DU0 mus musculu
21	4481.5	45.4	1527	MRP3_HUMAN	Q51433 homo sapien
22	4484.5	45.3	1533	059H05_HUMAN	Q59H05 homo sapien
23	4422	44.7	1514	096QA9_HUMAN	Q96QA9 homo sapien
24	4402	44.4	1498	059DK9_MOUSE	Q59DK9 mus musculu
25	4396	44.4	1522	MRP3_RAT	Q857R7 rattus norv
26	4148.5	41.9	1496	04STR7_TETNG	Q4STR7 tetraodon n
27	3885.5	39.2	1548	Q7KTC3_DROME	Q7KTC3 drosophila
28	3880.5	39.2	1548	Q7KTC1_DROME	Q7KTC1 drosophila
29	3879.5	39.2	1548	Q7KTC2_DROME	Q7KTC2 drosophila
30	3874.5	39.1	1548	Q7KTC2_DROME	Q7KTC2 drosophila
31	3860.5	39.0	1548	Q917N0_DROME	Q917N0 drosophila

ALIGNMENTS

RESULT 1	MRP1_HUMAN	STANDARD	PRT	1531 AA.
ID	P33527; 014819; P78419;			
AC	01-FEB-1994 (Rel. 28, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).			
GN	Name=ABCC1; Synonyms=MRP, MRP1;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RY	[1]			
RP	NUCLEOTIDE SEQUENCE. PubMed=1360704;			
RX	MEDLINE=93088080; PubMed=8098549;			
RA	Cole S.P.C., Bhargava J.G., Gerlach J.H., Mackie J.E., Grant C.E.,			
RA	Almquist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.,			
RT	"Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line."			
RL	Science 258:1650-1654(1992).			
RL	[2]			
RP	SEQUENCE REVISION. PubMed=8098549;			
RX	MEDLINE=93262415; PubMed=8098549;			
RA	Cole S.P.C., Deeley R.G.;			
RT	"Multidrug resistance-associated protein: sequence correction."			
RL	Science 260:879-879(1993).			
RL	[3]			
RP	NUCLEOTIDE SEQUENCE. PubMed=8649356;			
RX	MEDLINE=96251691; PubMed=8649356;			
RA	Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C.,			
RA	Deeley R.G.;			
RT	"Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter."			
RL	Mol. Pharmacol. 49:962-971(1996).			
RL	[4]			
RP	NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING. PubMed=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;			
RX	Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;			
RA	"Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA."			
RL	Genomics 45:368-378(1997).			
RL	[5]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA] OF 1131-1531. MEDLINE=99452270; PubMed=10493829; DOI=10.1006/geno.1999.5927;			
RX	Lothman B.J., Kim U.-J., Shedd V.P., Kalush F., Brandon R.,			
RA	Fuhrmann U., Mason T., Crosby M.L., Barnstead M., Cronin L.,			
RA	Edelman M., Harris P.C., Venter J.C., Adams M.D.;			
RT	"Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q."			
RL	Genomics 60:295-308(1999).			

32	3857	38.9	1549	2	Q7KTC0_DROME	Q7KTC0 drosophila
33	3845.5	38.8	1548	2	Q8T9C5_DROME	Q8T9C5 drosophila
34	3838.5	38.8	1548	2	Q7KTB8_DROME	Q7KTB8 drosophila
35	3836.5	38.7	1548	2	Q7KTB9_DROME	Q7KTB9 drosophila
36	3818.5	38.6	1548	2	Q7KTD0_DROME	Q7KTD0 drosophila
37	3813.5	38.5	1548	2	Q7KTC8_DROME	Q7KTC8 drosophila
38	3812.5	38.5	1548	2	Q7KTC4_DROME	Q7KTC4 drosophila
39	3807.5	38.4	1548	2	Q7KTC9_DROME	Q7KTC9 drosophila
40	3793.5	38.3	1548	2	Q7KTC7_DROME	Q7KTC7 drosophila
41	3790	38.3	1549	2	Q9VK56_DROME	Q9VK56 drosophila
42	3771.5	38.1	1548	2	Q7KTC5_DROME	Q7KTC5 drosophila
43	3769.5	38.1	1548	2	Q7KTC6_DROME	Q7KTC6 drosophila
44	3741	37.8	1564	2	Q8Q998_RAJER	Q8Q998 raja erinac
45	3738.5	37.8	1567	2	Q6PH26_BRARE	Q6PH26 brachydantio

[6] TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.
 RP MEDLINE=9744425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;
 RX Hifner D.R., Almquist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,
 RA Deley R.G., Cole S.P.C.;
 RT "Membrane topology of the multidrug resistance protein (MRP). A study
 of glycosylation-site mutants reveals an extracytosolic NH2
 terminus.";
 RL J. Biol. Chem. 272:23623-23630 (1997).
 RN [7]
 RP TOPOLOGY.
 RX MEDLINE=97476249; PubMed=9334225; DOI=10.1074/jbc.272.42.26479;
 RA Kaet C., Gros P.;
 RT "Topology mapping of the amino-terminal half of multidrug resistance-
 associated protein by epitope insertion and immunofluorescence.";
 RL J. Biol. Chem. 272:26479-26487 (1997).
 RN [8]
 RP TOPOLOGY.
 RX MEDLINE=96153110; PubMed=9485377; DOI=10.1021/bi972332v;
 RA Kaet C., Gros P.;
 RT "Epitope insertion favors a six transmembrane domain model for the
 carboxy-terminal portion of the multidrug resistance-associated
 protein.";
 RL Biochemistry 37:2305-2313 (1998).
 RN [9]
 RP MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-GLU-1455.
 RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;
 RA Cui L., Hou Y.-X., Riordan J.R., Chang X.-B.;
 RT "Mutations of the Walker B motif in the first nucleotide binding
 domain of multidrug resistance protein MRP1 prevent conformational
 maturation.";
 RL Arch. Biochem. Biophys. 392:153-161 (2001).
 RN [10]
 RP MUTAGENESIS OF TRP-1246.
 RX MEDLINE=21238219; PubMed=11278867; DOI=10.1074/jbc.M011246200;
 RA Ito K., Olsen S.L., Qiu W., Deley R.G., Cole S.P.C.;
 RT "Mutation of a single conserved tryptophan in multidrug resistance
 protein 1 (MRP1/ABCC1) results in loss of drug resistance and
 selective loss of organic anion transport.";
 RL J. Biol. Chem. 276:15616-15624 (2001).
 RN [11]
 RP VARIANTS GLN-633 AND VAL-671.
 RX MEDLINE=2029630; PubMed=10835642; DOI=10.1038/76102;
 RA Le Saux O., Urban Z., Tschuch C., Gaiszar K., Bacchelli B.,
 RA Quaglini D., Pasquali-Ronchetti I., Pope F.M., Richards A., Terry S.,
 RA Becroft L., de Paeppe A., Boyd C.D.;
 RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma
 elasticum.";
 RL Nat. Genet. 25:223-227 (2000).
 RN [12]
 RP VARIANT VAL-671.
 RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;
 RA Ringseil F., Lebowitz M.G., Cristiano A.M., Uitto J.;
 RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a
 transmembrane ATP-binding cassette (ABC) transporter.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006 (2000).
 RN [13]
 RP VARIANT SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.
 RX MEDLINE=21576494; PubMed=11721885; DOI=10.1007/s100380170017;
 RA Conrad S., Kaufmann H.-W., Ito K., Deley R.G., Cole S.P.C.,
 RA Schrenk D.;
 RT "Identification of human multidrug resistance protein 1 (MRP1)
 mutations and characterization of a G671V substitution.";
 RL J. Hum. Genet. 46:656-663 (2001).
 RN [14]
 RP VARIANTS MET-117 AND LEU-1512.
 RX MEDLINE=20579883; PubMed=11139250;
 DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14>3.0.CO;2-F;
 RA Perdut J., Germain D.P.;
 RT "Identification of novel polymorphisms in the pms and MRP1 (ABCC1)
 genes at locus 16p13.1 and exclusion of both genes as responsible for
 pseudoxanthoma elasticum.";
 RL Hum. Mutat. 17:74-75 (2001).

[15]
 RP VARIANTS SER-433, ILE-73; GLN-723 AND GLN-1058.
 RX MEDLINE=21163848; PubMed=11266082;
 DOI=10.1097/00008571-200103000-00008;
 RA Ito S., Ito I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;
 RT "Polymorphism of the ABC transporter genes, MRP1, MRP2 and MRP2/cMOAT,
 in healthy Japanese subjects.";
 RL Pharmacogenetics 11:175-184 (2001).
 CC -1- FUNCTION: May participate directly in the active transport of
 drugs into subcellular organelles or influence drug distribution
 indirectly.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=8;
 CC Comment=Additional isoforms seem to exist. Experimental
 CC confirmation may be lacking for some isoforms;
 CC Name=All isoforms;
 CC IsoId=P33527-1; Sequence=Displayed;
 CC Name=Delexon-17;
 CC IsoId=P33527-2; Sequence=VSP_000037;
 CC Name=Delexon-18;
 CC IsoId=P33527-3; Sequence=VSP_000038;
 CC Name=Delexon-30;
 CC IsoId=P33527-4; Sequence=VSP_000039;
 CC Name=Delexon-17-18;
 CC IsoId=P33527-5; Sequence=VSP_000037, VSP_000038;
 CC Name=Delexon-17-30;
 CC IsoId=P33527-6; Sequence=VSP_000037, VSP_000039;
 CC Name=Delexon-18-30;
 CC IsoId=P33527-7; Sequence=VSP_000038, VSP_000039;
 CC Name=Delexon-17-18-30;
 CC IsoId=P33527-8; Sequence=VSP_000037, VSP_000038, VSP_000039;
 CC -1- TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear
 CC cells.
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
 CC -1- SIMILARITY: Contains 2 ABC transporter domains.
 CC -1- DATABASE: NME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.inbiohogen.fr/services/chromocancer/genes/MRP106.html".
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: L05628; AAB4616.1; -, mRNA.
 CC EMBL: AF022853; AAB83983.1; -, Genomic DNA.
 CC EMBL: AF022824; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022825; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022826; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022827; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022828; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022829; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022830; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022831; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022832; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022833; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022834; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022835; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022836; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022837; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022838; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022839; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022840; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022841; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022842; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022843; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022844; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022845; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022846; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022847; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022848; AAB83983.1; JOINED; Genomic DNA.

DR EMBL; AF022849; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83983.1; JOINED; Genomic DNA.

Query Match 79.4%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALGFCGADSDPFLMDNMTNTSNPDTCFQNTVWVPCFPLMACPPFFLYLSRH 60
DB 1 MALGFCGADSDPFLMDNMTNTSNPDTCFQNTVWVPCFPLMACPPFFLYLSRH 60

QY 61 DRGYIOMTPLNKTATLGLFLLIMVCMADLFYSFWMRSRGIFLAPVFLPILLGITLLA 120
DB 61 DRGYIOMTPLNKTATLGLFLLIMVCMADLFYSFWMRSRGIFLAPVFLPILLGITLLA 120

QY 121 TFLIQLERKGVOSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLFRDITFYVFS 180
DB 121 TFLIQLERKGVOSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLFRDITFYVFS 180

QY 181 LLLIQLVLSGSDSPLESETIHDNPPCESSASFLSITFMWITGLIVRGYROPLEGGSD 240
DB 181 LLLIQLVLSGSDSPLESETIHDNPPCESSASFLSITFMWITGLIVRGYROPLEGGSD 240

QY 241 LMSLNKEDTSBOVVPVLVKNMKKCAKTRKQPVKVYSSKDPAOPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSBOVVPVLVKNMKKCAKTRKQPVKVYSSKDPAOPKSSKYDANEVEAL 300

QY 301 IVKSPPOKEMNPSLRKVLYKTPGYFLMSFPFKAHDLMSFGPOLIKLIFVNDTKAD 360
DB 301 IVKSPPOKEMNPSLRKVLYKTPGYFLMSFPFKAHDLMSFGPOLIKLIFVNDTKAD 360

QY 361 WQGFYVTLFVTACLOTLVTHOYFHCFSVGMRTKTAIVGAVRKALVITNSARKSTV 420
DB 361 WQGFYVTLFVTACLOTLVTHOYFHCFSVGMRTKTAIVGAVRKALVITNSARKSTV 420

QY 421 GEIYNLMSVDAQREMDLATYINMIWSAPLOYIALYLLMLAGPSVLAGVAVMLAYVN 480
DB 421 GEIYNLMSVDAQREMDLATYINMIWSAPLOYIALYLLMLAGPSVLAGVAVMLAYVN 480

QY 481 AVMAKTKTYOVAHMKSKDNRIKLMNEILNGIKYKLAYAMELAFKDYALIROBELKYLK 540
DB 481 AVMAKTKTYOVAHMKSKDNRIKLMNEILNGIKYKLAYAMELAFKDYALIROBELKYLK 540

QY 541 KSAVLAVGTFTWCTPPLVALCTPAYVVTIDENNIIDAOAFVSLFNLIRPPLILP 600
DB 541 KSAVLAVGTFTWCTPPLVALCTPAYVVTIDENNIIDAOAFVSLFNLIRPPLILP 600

QY 601 MVISSIVQASVSLKRLRIFLSHEELPDSIRRRPVKOGGNSITVRNATPTMARSDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELPDSIRRRPVKOGGNSITVRNATPTMARSDPT 660

QY 661 LMGITFSIPEGALVAVVGQVCGKSSLSALLAEMDKVEGVAIKGSVAVVPQQAAMTQND 720
DB 661 LMGITFSIPEGALVAVVGQVCGKSSLSALLAEMDKVEGVAIKGSVAVVPQQAAMTQND 720

QY 721 SLRNNILFCQOLEPPYRSVTOACALLPDLIILHSGDTEIGEGVNLSSGQKRVSLAR 780
DB 721 SLRNNILFCQOLEPPYRSVTOACALLPDLIILHSGDTEIGEGVNLSSGQKRVSLAR 780

QY 781 AVYSNADITYLPDDPLSAVDHVGKIFENVIGPKMLKNKRIIVTSMSTLPDQVDIIV 840
DB 781 AVYSNADITYLPDDPLSAVDHVGKIFENVIGPKMLKNKRIIVTSMSTLPDQVDIIV 840

QY 841 MSGGKISEMSYQELLARDGAFAPLRTYASTEOBDAEENGVTGSPGKEAKOMENGM 900
DB 841 MSGGKISEMSYQELLARDGAFAPLRTYASTEOBDAEENGVTGSPGKEAKOMENGM 900

QY 901 LYTDSAKKOLOROLSSSSSSGDISRHNSGTAELQKAEKEEFWKLEADKATGQYKL 960
DB 901 LYTDSAKKOLOROLSSSSSSGDISRHNSGTAELQKAEKEEFWKLEADKATGQYKL 960

QY 961 SYVYDYMKAIGLFLISFLIFLMCNHVSALASNYMLSLMTDPIVNGTOEHTKRLSVYG 1020
DB 961 SYVYDYMKAIGLFLISFLIFLMCNHVSALASNYMLSLMTDPIVNGTOEHTKRLSVYG 1020

DR 961 SYVYDYMKAIGLFLISFLIFLMCNHVSALASNYMLSLMTDPIVNGTOEHTKRLSVYG 1020

QY 1021 ALGISOGIAYRGYMAVSIIGIILASRCLHDHLSILRSPMSFFERPSSGLVNRFSREL 1080
DB 1021 ALGISOGIAYRGYMAVSIIGIILASRCLHDHLSILRSPMSFFERPSSGLVNRFSREL 1080

QY 1081 DTVDMSIPEVIKMGMSLFENVIGACIVILATPAAIIIPPLGLIYFVQRFYVASSROL 1140
DB 1081 DTVDMSIPEVIKMGMSLFENVIGACIVILATPAAIIIPPLGLIYFVQRFYVASSROL 1140

QY 1141 KRLESVSRPYSHFNETLGVSVIRAFEEDEERIHOSDLKYDENOKAYPSIYANRWLA 1200
DB 1141 KRLESVSRPYSHFNETLGVSVIRAFEEDEERIHOSDLKYDENOKAYPSIYANRWLA 1200

QY 1201 VRLCEVGNCTVLPALPAVVISRHSLSGLVGLSVYSIQVTTYANLVYRMSSEMTNIVA 1260
DB 1201 VRLCEVGNCTVLPALPAVVISRHSLSGLVGLSVYSIQVTTYANLVYRMSSEMTNIVA 1260

QY 1261 VERLKEYSETEKEAPMOIOETAPSSMPQVGRVEFRNYCLARYEDLPVLHINVTNGG 1320
DB 1261 VERLKEYSETEKEAPMOIOETAPSSMPQVGRVEFRNYCLARYEDLPVLHINVTNGG 1320

QY 1321 EKVGIVERTGAGKSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPDPVLF 1380
DB 1321 EKVGIVERTGAGKSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPDPVLF 1380

QY 1381 SGSLRNMLDPPSOYSDEBWTSLELAHLKOFVSLPKLHDECEGGENLSVGRQVCL 1440
DB 1381 SGSLRNMLDPPSOYSDEBWTSLELAHLKOFVSLPKLHDECEGGENLSVGRQVCL 1440

QY 1441 ARALLRRTKTLIVDEAPAAVLETDLIQSTIRQFEDCTVLTAAHLNITMDTRVYL 1500
DB 1441 ARALLRRTKTLIVDEAPAAVLETDLIQSTIRQFEDCTVLTAAHLNITMDTRVYL 1500

QY 1501 DKGEIOEGYGAAPSDLIQORGLFYSSAKDAGLV 1531
DB 1501 DKGEIOEGYGAAPSDLIQORGLFYSSAKDAGLV 1531

RESULT 2
09U099 HUMAN
ID 09U099 HUMAN PRELIMINARY; .PRT; 1515 AA.
AC 09U099;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2004 (Tremblrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name-MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=38008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz B.U., Cole S.P.C., Dealey R.G.;
RT "Analysis of the Intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF022853; AAB83979.1; -; Genomic DNA.
DR EMBL; AF022857; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022858; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022862; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022863; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022863; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022865; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022839; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022849; AAB83979.1; JOINED; Genomic DNA.

DR EMBL: AF022848; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022847; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022846; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022845; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022844; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022843; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022842; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022840; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022838; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022836; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022834; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022832; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022826; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022825; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022824; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022830; AAB83979.1; JOINED; Genomic DNA.
DR HSSP: P08716; IMT0.
DR Ensembl: ENSG00000103222; Homo sapiens.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR011527; ABC_membrane_1.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR005292; MRP_assoc.
DR InterPro: IPR00719; Prot_kinase.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; Abc_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS00929; ABC_TM1F; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR ATP-binding; Nucleotide-binding.
DR NON_TER 1
SQ SEQUENCE 1515 AA; 169853 MW; DC85592817C439FE CRC64;

Query Match 78.5%; Score 7769; DB 2; Length 1515;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DMNTWNTSNDFTCFQNTVTVWVPCFYLMACPFYFLYSRHDGTYOMTPLNKTKTA 76
DB 1 DMNTWNTSNDFTCFQNTVTVWVPCFYLMACPFYFLYSRHDGTYOMTPLNKTKTA 60
QY 77 LGFLIMVCMADLFYSFWERSRGIFLAPFLVSPFLIGITTLATFLIQLERRKVGSSG 136
DB 61 LGFLIMVCMADLFYSFWERSRGIFLAPFLVSPFLIGITTLATFLIQLERRKVGSSG 120
QY 137 IMLTFWIALVCAALILRSKIMTALKEDAQVDLFPDIFFYVFSLLQLVYSCFSDDSP 196
DB 121 IMLTFWIALVCAALILRSKIMTALKEDAQVDLFPDIFFYVFSLLQLVYSCFSDDSP 180
QY 197 LFSETHDPNCPRESSASFSLRTFMWITGLIVRGYROPLEGSDMLSKEDPESQVAV 256
DB 181 LFSETHDPNCPRESSASFSLRTFMWITGLIVRGYROPLEGSDMLSKEDPESQVAV 240
QY 257 LVKMKKCAKTRKQPVKVYVSSKDPAPQKSSKVDANEVEVALIVKSPQKEMNSLPKV 316
DB 241 LVKMKKCAKTRKQPVKVYVSSKDPAPQKSSKVDANEVEVALIVKSPQKEMNSLPKV 300
QY 317 LYKTGPFPLMSFFPKAIIHDLMMFSGPOLIKLIRVANDTKAPDWQGYTYTTLVTVACL 376
DB 301 LYKTGPFPLMSFFPKAIIHDLMMFSGPOLIKLIRVANDTKAPDWQGYTYTTLVTVACL 360

QY 377 QTVLHQYFHI CFSVGNRIKTAIVIGAVYRKALVITNSARKSITGEIVNLSMDAQRPM 436
DB 361 QTVLHQYFHI CFSVGNRIKTAIVIGAVYRKALVITNSARKSITGEIVNLSMDAQRPM 420
QY 437 LATYINNIWGNPQVLIATLILMTNIGPSYLAAGVAVNWVAVNVAWAMKTKTYQVAMK 496
DB 421 LATYINNIWGNPQVLIATLILMTNIGPSYLAAGVAVNWVAVNVAWAMKTKTYQVAMK 480
QY 497 SKDNRIKLMNEILNGIKYLKYLAMELAFKDKVLAIRQEBELVKLSAYLSAVGTFTWCT 556
DB 481 SKDNRIKLMNEILNGIKYLKYLAMELAFKDKVLAIRQEBELVKLSAYLSAVGTFTWCT 540
QY 557 PELVALCTFAVYVITDENNIIDAQTAVSALFNILRPPLNIPMWISSIVQASVSLKRL 616
DB 541 PELVALCTFAVYVITDENNIIDAQTAVSALFNILRPPLNIPMWISSIVQASVSLKRL 600
QY 617 RIFLSHELEPDSIERPVDGGGTNSITVRNAFTMARSPPTLNGITPSIRPGALVAV 676
DB 601 RIFLSHELEPDSIERPVDGGGTNSITVRNAFTMARSPPTLNGITPSIRPGALVAV 660
QY 677 VGQVCGCKSSLALAEMLDVEGVHAIKGSVAIVPOQAWIQNDLSRENILFGQLEBPY 736
DB 661 VGQVCGCKSSLALAEMLDVEGVHAIKGSVAIVPOQAWIQNDLSRENILFGQLEBPY 720
QY 737 YRSVIQACALLPDLEILPSGDRTEIGEGVNLSCGQKQKVSABAVYSNADIVLPDPLS 796
DB 721 YRSVIQACALLPDLEILPSGDRTEIGEGVNLSCGQKQKVSABAVYSNADIVLPDPLS 780
QY 797 AVDAHVGHIIFENYIGPKMLKNKTRILLYTHSMSTYLPQVDYIYMSGKXISEMSYDGL 856
DB 781 AVDAHVGHIIFENYIGPKMLKNKTRILLYTHSMSTYLPQVDYIYMSGKXISEMSYDGL 840
QY 857 ARDGAFAFLRTYASTEODPAEENGVTGVSQPKGEAKOMENGLVTDGAKQIQORLSS 916
DB 841 ARDGAFAFLRTYASTEODPAEENGVTGVSQPKGEAKOMENGLVTDGAKQIQORLSS 900
QY 917 SSSYSGDISRRHNSTAELQKAEAKKETWKLMEADKAQTGVKLSVYWDYKAIGLFTSF 976
DB 901 SSSYSGDISRRHNSTAELQKAEAKKETWKLMEADKAQTGVKLSVYWDYKAIGLFTSF 960
QY 977 LSTIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTKRLSVGALGISQIIAVPGYMA 1036
DB 961 LSTIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTKRLSVGALGISQIIAVPGYMA 1020
QY 1037 VSIGGIIASRCLAHVDLHSLIRSPMSFPERTPSGNLVNRFSKELDYDMSIPEVYKPMFG 1096
DB 1021 VSIGGIIASRCLAHVDLHSLIRSPMSFPERTPSGNLVNRFSKELDYDMSIPEVYKPMFG 1080
QY 1097 SLFNVIGACIVILLATPIAIIIPPLGLIYEFVQRFYVASSRQLKRLSVSRSPYSHFN 1156
DB 1081 SLFNVIGACIVILLATPIAIIIPPLGLIYEFVQRFYVASSRQLKRLSVSRSPYSHFN 1140
QY 1157 ETLIGVSTIRAFEOERFIHOSDKVDENQAAVPSIYANRWLAVRLECNGENCIYLPAL 1216
DB 1141 ETLIGVSTIRAFEOERFIHOSDKVDENQAAVPSIYANRWLAVRLECNGENCIYLPAL 1200
QY 1217 FAVISRHSLSAGLVGSYSLOYTTYTNMIVRMSSEMETNIIVAVRLKEYSSETEKAPW 1276
DB 1201 FAVISRHSLSAGLVGSYSLOYTTYTNMIVRMSSEMETNIIVAVRLKEYSSETEKAPW 1260
QY 1277 QIOETAPPSWPOVGRVFRNYCLRYREDLDFVLRIHNVITINGEKVGI VGRTAGKSSL 1336
DB 1261 QIOETAPPSWPOVGRVFRNYCLRYREDLDFVLRIHNVITINGEKVGI VGRTAGKSSL 1320
QY 1337 TLGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDVLPSGSLRNMIDPSSQYSD 1396
DB 1321 TLGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDVLPSGSLRNMIDPSSQYSD 1380
QY 1397 EEWMTSLELAHKQFVSALPDKLDHECABEGENISVQGRQIVCLARALLRKTKIIVLBEA 1456
DB 1381 EEWMTSLELAHKQFVSALPDKLDHECABEGENISVQGRQIVCLARALLRKTKIIVLBEA 1440
QY 1457 TAAVDLLETDDLIOSTIRTOFEDCTVLTIARHLNLTIMDYTRVIVLDKGEIOEYGA PSDLLQ 1516

Db 1441 TAAVLETDLIQSTIFRQFEDCTVLTIAHLRLNTMDTRVYLDKGEIQEIGAPSDLLQ 1500
QY 1517 ORGLFYSAKXAGLV 1531
Db 1501 QRGIFYSAKXAGLV 1515

RESULT 3

0864R9 MACPA PRELIMINARY; PRT; 1531 AA.
ID 0864R9; AC 0864R9;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Multidrug resistance protein 1B.
GN Name=Mkpl.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecidae; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22544876; PubMed=12657726;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RT Dantzig A.H., Perry W.L.,
RT "Cloning and functional characterization of the multidrug resistance-
RT associated protein (MRP1/ABCC1) from the cynomolgus monkey";
RN [2]
RP NUCLEOTIDE SEQUENCE.

RA Perry W.L. III, Godinot N.; EMBL/GenBank/DBJ databases.
RL Submitted (Aug-2002) to the ABC transporter family.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY146673; AAN65349.1; -; mRNA.
DR HSBP; P08716; IMTO.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042666; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR Prodom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS0107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1531 AA; 171659 MW; 1AE788EFDF9EF459 CRC64;

Query Match 77.7%; Score 7695; DB 2; Length 1531;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLDMNMTNTSNPDTKCFQNTVLMVWPCFYLMACFPFYLYSRH 60
Db 1 MALRGFCSADGSDPLDMNMTNTSNPDTKCFQNTVLMVWPCFYLMACFPFYLYSRH 60
QY 61 DRGVIQMTPLNKTALGFLIMVGMADLFYSFMEGRGIFLADVLVSPFLIGITTLA 120
Db 61 DRGVIQMTPLNKTALGFLIMVGMADLFYSFMEGRGIFLADVLVSPFLIGITTLA 120

QY 121 TFLIQLERRKGVSSGIMLTFTMLVALVCALALISKIMTALKEADQVDFPDITFYVES 180
Db 121 TFLIQLERRKGVSSGIMLTFTMLVALVCALALISKIMTALKEADQVDFPDITFYVES 180
QY 181 LLLIQVLVSCPSDRSPLFSETHIDPNCPRESSASFLSRITFTWTTGLIVRGYRPLBESD 240
Db 181 LLLIQVLVSCPSDRSPLFSETHIDPNCPRESSASFLSRITFTWTTGLIVRGYRPLBESD 240
QY 241 LMSLNKEDTSQVIVLVKMKKCACTRKQPVYVYSSKQPAQKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSQVIVLVKMKKCACTRKQPVYVYSSKQPAQKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLKVLYKTEGPGYFLMSFFPKALHDLMMFSGPEILKLLINFDVDTKPD 360
Db 301 IVKSPQKEMNPSLKVLYKTEGPGYFLMSFFPKALHDLMMFSGPEILKLLINFDVDTKPD 360
QY 361 WQGYFYVLLFVTACLOTVLVHQYFHLCPVSGMKIKTAVIGAVRKALVITNSARKSSTV 420
Db 361 WQGYFYVLLFVTACLOTVLVHQYFHLCPVSGMKIKTAVIGAVRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRPMDLATYINMTWSAPLOVITALLYMLNGPSVLAGAVMYLVNPN 480
Db 421 GEIVNLSVDAQRPMDLATYINMTWSAPLOVITALLYMLNGPSVLAGAVMYLVNPN 480
QY 481 AVMAKTKTQVAMKSKDNRIKLMNELLNGIKYLKIYAMELAKDKVLATROBELKYLK 540
Db 481 AVMAKTKTQVAMKSKDNRIKLMNELLNGIKYLKIYAMELAKDKVLATROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIDDAQAFSLALFNILRPPLNLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIDDAQAFSLALFNILRPPLNLP 600
QY 601 MTISSIVQASVSLKRLIFLSHEELPDSIERRPVKDGGTNSITVNAATFTMARSDPT 660
Db 601 MTISSIVQASVSLKRLIFLSHEELPDSIERRPVKDGGTNSITVNAATFTMARSDPT 660
QY 661 LNTGTFIPRGALVAVVGVCGCGKSLLSALLAEMDVEGHVALKGSVAVYPOQAMOND 720
Db 661 LNTGTFIPRGALVAVVGVCGCGKSLLSALLAEMDVEGHVALKGSVAVYPOQAMOND 720
QY 721 SLRENILFGQLEBPYRSYIACALPDLIELPSGRTEIGEKNVLSGGQKQVSLAR 780
Db 721 SLRENILFGQLEBPYRSYIACALPDLIELPSGRTEIGEKNVLSGGQKQVSLAR 780
QY 781 AVYSNADIVYFDPLSAVDAAVGHKIFENVYIGPKMKLNKTRILLVTHSMYSYLPQVDYIY 840
Db 781 AVYSNADIVYFDPLSAVDAAVGHKIFENVYIGPKMKLNKTRILLVTHSMYSYLPQVDYIY 840
QY 841 MSGGKISEMGSYQELLRDGAFAEFLRTYASAEQODPEINGVYGVSGPKKAKQMNEM 900
Db 841 MSGGKISEMGSYQELLRDGAFAEFLRTYASAEQODPEINGVYGVSGPKKAKQMNEM 900
QY 901 LVYDSAGKQROLORSSSSYSYSDISRHHNSTAELOKAEKKEETKMLEADKQOTGOVKL 960
Db 901 LVYDSAGKQROLORSSSSYSYSDISRHHNSTAELOKAEKKEETKMLEADKQOTGOVKL 960
QY 961 SVYWDYMKALGFLISFLISFLFMCNHYASALASNYWLSLWTDPIVNGTQEHTRVLSYG 1020
Db 961 SVYWDYMKALGFLISFLISFLFMCNHYASALASNYWLSLWTDPIVNGTQEHTRVLSYG 1020
QY 1021 ALGISQGIAPFGYSMAVSTIGIILASRCLAHVDLHSILRSMSFPERPSGNLVNRSKEL 1080
Db 1021 ALGISQGIAPFGYSMAVSTIGIILASRCLAHVDLHSILRSMSFPERPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPVYIKMFMSLFNVIGACTYIILATPLAIIIPPLGLIYFVORFYVASSROL 1140
Db 1081 DTVDSMIPVYIKMFMSLFNVIGACTYIILATPLAIIIPPLGLIYFVORFYVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLLIGSVIRAFESQERPIHQSDIKVDENQKAYYPSIIVANWLA 1200
Db 1141 KRLSVSRSPVYSHFNETLLIGSVIRAFESQERPIHQSDIKVDENQKAYYPSIIVANWLA 1200
QY 1201 VRLCEVGNCLVPAALFAVISRSLSAGLVGSLVSLOVTTYLLNMLVRRSSMEHTNIVA 1260

Db 1201 VRLCEVGCICVLFALFAVIRSHSISAGLVLSVYSIQVTTYINMLVMSSEMETNIVA 1260
Qy 1261 VERIKESSEFEKEAPMOIQETAPSPMPOGVERBANCAYREDLDVLRHINVTNGG 1320
Db 1261 VERIKESSEFEKEAPMOIQETAPSPMPOGVERBANCAYREDLDVLRHINVTNGG 1320
Qy 1321 EKVGI VGRGTAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITTIPQDPVLF 1380
Db 1321 EKVGI VGRGTAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITTIPQDPVLF 1380
Qy 1381 SGIIRPMNIDPFSQYSDSEEWTSLELAHLKDFVSALPDYLDHCAEGENLSVGORQVCL 1440
Db 1381 SGIIRPMNIDPFSQYSDSEEWTSLELAHLKDFVSALPDYLDHCAEGENLSVGORQVCL 1440
Qy 1441 ARALLRKTIIVLBATAVLETDLLQSTIRIQFOEDCTVLTAAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKTIIVLBATAVLETDLLQSTIRIQFOEDCTVLTAAHRLNTIMDYTRVIVL 1500
Qy 1501 DKGEIQEYGAPSDLLQGRGLFYMAKDAGLV 1531
Db 1501 DKGEIQEYGAPSDLLQGRGLFYMAKDAGLV 1531

RESULT 4

Q864S0 MACFA PRELIMINARY; PRT; 1531 AA.

AC Q864S0
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein 1A.
GN Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopitheciidae; Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22544876; PubMed=12657726;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Perry W.L. III, Godinot N.;
RL Submitted (AUG-2002) to the EMBL/Genbank/DBJ databases.
CC -1- STRIKELARITY: belongs to the ABC transporter family.
EMBL: AY146672; AAN65348.1; -, mRNA.
DR HSSP; P08716; 1MT0.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGFams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.

DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1531 AA, 171701 MW, 16871286CC2D2B89 CRC64;

Query Match 77.6%; Score 7689; DB 2; Length 1531;

Best Local Similarity 97.6%; Pred. No. 0;
Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MALRGFSGADGSDPLMDMNTWNTSNPDFTKCPONTLVWPCRYLVACPFYFLYLSRH 60
Db 1 MALRGFSGADGSDPLMDMNTWNTSNPDFTKCPONTLVWPCRYLVACPFYFLYLSRH 60
Qy 61 DRGYIOMTPUNKKTALGFLLMIVCMADLFYSFERSRGIFLAIVFLVSPFLIGITLLA 120
Db 61 DRGYIOMTPUNKKTALGFLLMIVCMADLFYSFERSRGIFLAIVFLVSPFLIGITLLA 120
Qy 121 TFLIQLEBRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180
Db 121 TFLIQLEBRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180
Qy 121 TFLIQLEBRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180
Db 121 TFLIQLEBRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180
Qy 181 LLLIQVLSCFSDSPLEFSETIHDNCPRESSASFLSRTIFWMTTGLIVRGYRQPLBSSD 240
Db 181 LLLIQVLSCFSDSPLEFSETIHDNCPRESSASFLSRTIFWMTTGLIVRGYRQPLBSSD 240
Qy 241 LMSLNKEDTSEQVAVPVLYKMKKECAKTRKQPVKVVYSSKDPAOKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVAVPVLYKMKKECAKTRKQPVKVVYSSKDPAOKESSKYDANEVEAL 300
Qy 301 IVKSPQKEMNPSLFRVLYKTGFPYFLMSFFPKAIDHLMFSGPOILKILIFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFRVLYKTGFPYFLMSFFPKAIDHLMFSGPOILKILIFVNDTKAPD 360
Qy 361 WQGFYTYTLVATACLOTVLHOYFHI CFVSGMIRKTRAVIAVVRKALVITNAKRSSTV 420
Db 361 WQGFYTYTLVATACLOTVLHOYFHI CFVSGMIRKTRAVIAVVRKALVITNAKRSSTV 420
Qy 421 GEIVNLSVDAQRFPMDLATYINMISAPLQVTLALYLMLNLGSPVLAVGVAVMYLVN 480
Db 421 GEIVNLSVDAQRFPMDLATYINMISAPLQVTLALYLMLNLGSPVLAVGVAVMYLVN 480
Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNBIINGIKVLYKLYAMELAFKDYALAIROBELKYLK 540
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNBIINGIKVLYKLYAMELAFKDYALAIROBELKYLK 540
Qy 541 KSAIYLSANGTFTWCTPFLVALCTPAVYVTTIDENNIIIDAQAFAPSLAFNLRPPLNLP 600
Db 541 KSAIYLSANGTFTWCTPFLVALCTPAVYVTTIDENNIIIDAQAFAPSLAFNLRPPLNLP 600
Qy 601 MYISIVQASVSLKRLRIFLSHEELEPDSIRRRPVKDGCTNSITVRNATFTWASDPPT 660
Db 601 MYISIVQASVSLKRLRIFLSHEELEPDSIRRRPVKDGCTNSITVRNATFTWASDPPT 660
Qy 661 LINGITFSIPREGALVAVVQVCGKSSLLSALLAEWDKVEGHVAIKGSVAVVYVPOQAMIOND 720
Db 661 LINGITFSIPREGALVAVVQVCGKSSLLSALLAEWDKVEGHVAIKGSVAVVYVPOQAMIOND 720
Qy 721 SLRENTILEGCGLEEPYRHSVYQAALLPDLLELRSGBDTEI GEGKVNLSGGQKQVSLAR 780
Db 721 SLRENTILEGCGLEEPYRHSVYQAALLPDLLELRSGBDTEI GEGKVNLSGGQKQVSLAR 780
Qy 781 AVYSNADYILPDDPLSAVDAHVKHIFENVIGPKMLKNKTRILVTHSMYSYLPQVDVIV 840
Db 781 AVYSNADYILPDDPLSAVDAHVKHIFENVIGPKMLKNKTRILVTHSMYSYLPQVDVIV 840
Qy 841 MSGKISMGSYOEVLARDGAFAEFLRTYASTEDQDAENGVTYVSGPGKAKEMKMG 900
Db 841 MSGKISMGSYOEVLARDGAFAEFLRTYASTEDQDAENGVTYVSGPGKAKEMKMG 900
Qy 901 LVYDSAGKOLQOLSSSSSYSGDI SRHNSTAELOKAEKKETTKLMEADKAOQGOYKL 960
Db 901 LVYDSAGKOLQOLSSSSSYSGDI SRHNSTAELOKAEKKETTKLMEADKAOQGOYKL 960

QY 961 SVYDYMKAIGLFTISFLISIFLEMCNHSALASNYWLSMTDDPTVNGTQEHKRLSYVG 1020
DB 961 SVYDYMKAIGLFTISFLISIFLEMCNHSALASNYWLSMTDDPTVNGTQEHKRLSYVG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGGILASRYLHVLDLHSLSPMSFFERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGGILASRYLHVLDLHSLSPMSFFERTPSGNLVNRFSEKL 1080
QY 1081 DTYSMTPEVTKMFMGSLFNVIACIVILATPTIAIIPPLGLIYFVQRFYVASSQOL 1140
DB 1081 DTYSMTPEVTKMFMGSLFNVIACIVILATPTIAIIPPLGLIYFVQRFYVASSQOL 1140
QY 1141 KRLSVSRSPVYSHFNETLIGSVIRAFEBQERFIHOSDLVYDENOKAYYSIVANRMLA 1200
DB 1141 KRLSVSRSPVYSHFNETLIGSVIRAFEBQERFIHOSDLVYDENOKAYYSIVANRMLA 1200
QY 1201 VRLFCVGNICIVLPAALPAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRRMSSEMETNIVA 1260
DB 1201 VRLFCVGNICIVLPAALPAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIQETAPPSNPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPMOIQETAPPSNPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVGIVERTGAGKSLTLGLFRINESAGEIITIDGINAKTGLADAPFKITIIPODPYLF 1380
DB 1321 EKVGIVERTGAGKSLTLGLFRINESAGEIITIDGINAKTGLADAPFKITIIPODPYLF 1380
QY 1381 SGSLRNMJLDPFSQYSDSEVWTSLELAHLKDFVSALPKDLHBCAEGENLSVQROVLCL 1440
DB 1381 SGSLRNMJLDPFSQYSDSEVWTSLELAHLKDFVSALPKDLHBCAEGENLSVQROVLCL 1440
QY 1441 ARALLRTKTLIVDEATPAVDLETDLLIQSTIRFQEDCTVLTAAHRLNTIMDYTRVYL 1500
DB 1441 ARALLRTKTLIVDEATPAVDLETDLLIQSTIRFQEDCTVLTAAHRLNTIMDYTRVYL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYMAKADGLV 1531
DB 1501 DKGEIOEYGA PSDLLQORGLFYMAKADGLV 1531

RESULT 5
Q9U097_HUMAN
ID Q9U097_HUMAN PRELIMINARY; PRT; 1459 AA.
AC Q9U097;
DT 01-MAY-2000 (Tremblrel, 13, Created)
DT 01-MAY-2000 (Tremblrel, 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel, 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name-MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-376(1997).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF022838; AAB83980.1; -; Genomic DNA.
DR EMBL; AF022834; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022824; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022826; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022828; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83980.1; JOINED; Genomic DNA.

DR EMBL; AF022838; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022848; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022839; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022852; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022851; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022829; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022827; AAB83980.1; JOINED; Genomic DNA.
DR HSSP; P08716; 1MT0.
DR Ensemble; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transp; 1.
DR SMART; SM00382; AAA; 2.
DR TIGRfam; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PSS0929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PSS0893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KM ATP-binding, Nucleotide-binding.
FT NON_TER 1
SQ SEQUENCE 1459 AA; 16381 MW; 6A016A5A6E7A95C CRC64;
Query Match 75.3%; Score 7444; DB 2; Length 1459;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1459; Conservative 0; Mismatches 0; Indels 56; Gaps 1;
QY 17 DNVVTNNTSNPDPTKCFQNTLVVWVPCFYLMACFPFYFLYSRHDGTYIOMTPIANKTKTA 76
DB 1 DNVVTNNTSNPDPTKCFQNTLVVWVPCFYLMACFPFYFLYSRHDGTYIOMTPIANKTKTA 76
QY 77 LGFLNIYVCADLIFYSWERSRGIFLAPVFLVSPPTLLGTTTLATFLIQLERRKGVQSSG 136
DB 61 LGFLNIYVCADLIFYSWERSRGIFLAPVFLVSPPTLLGTTTLATFLIQLERRKGVQSSG 120
QY 137 IMTLFNLVNLVLCALATIRSKIMFALKEDQVDFRIDTFVYVSSLLLIOLVLSGFSRSP 196
DB 121 IMTLFNLVNLVLCALATIRSKIMFALKEDQVDFRIDTFVYVSSLLLIOLVLSGFSRSP 180
QY 197 LFSETHDPNCPRESSASFLSRITFWITGLIVRGYROPLEGSDLSLNKEDTSEQVVPV 256
DB 181 LFSETHDPNCPRESSASFLSRITFWITGLIVRGYROPLEGSDLSLNKEDTSEQVVPV 240
QY 257 LVNWKKECKTKTKOPKVVYSSKDPAPQKESKVDANNEVELIYKSPKKNPSPISFKV 316
DB 241 LVNWKKECKTKTKOPKVVYSSKDPAPQKESKVDANNEVELIYKSPKKNPSPISFKV 300
QY 317 LVYTFGPYFAMSFPAIHIDMMFSGPOLIKLIKVNTPDKADPMQGYFTYVTLPTVACL 376

Db 301 LYKTEGPEYELMSPEFFKAIHIDIMFSGPOLIKLIXEVNDTAPDMOGYFYTVLLFVTACL 360
Qy 377 QTVLHVQYFHIQFVSGNRIKTAIVIGAVYRKALVITNSARKSSTVGEIYNLMSVDAQRPMD 436
Db 361 QTVLHVQYFHIQFVSGNRIKTAIVIGAVYRKALVITNSARKSSTVGEIYNLMSVDAQRPMD 420
Qy 437 LATYINMISAPLOVITLALYILMLNIGPSVLAGVAVMTLVMPVNAVMAMKTKTYQVAHMK 496
Db 421 LATYINMISAPLOVITLALYILMLNIGPSVLAGVAVMTLVMPVNAVMAMKTKTYQVAHMK 480
Qy 497 SKDNRIKIMNEILINGIKYLKLYAMELAKDKVIALROBELKYLKKSATLSAVGFTWCT 556
Db 481 SKDNRIKIMNEILINGIKYLKLYAMELAKDKVIALROBELKYLKKSATLSAVGFTWCT 540
Qy 557 PFLVALCTFAVYVITIDENNIIDAOQAFVSLALFNILRPPLNILLPWVSIIVQASVSLKRL 616
Db 541 PFLVALCTFAVYVITIDENNIIDAOQAFVSLALFNILRPPLNILLPWVSIIVQASVSLKRL 600
Qy 617 RIPLSHEELPDSIERRPVKOGGTSITVRNATFTMARSDPPLNLTFTPSIPGALVAV 676
Db 601 RIPLSHEELPDSIERRPVKOGGTSITVRNATFTMARSDPPLNLTFTPSIPGALVAV 660
Qy 677 VGVGCGGSSLSLALAMDQVEGHVATKGSVAVYPOQAMTQNDLSRENILLPGCOLBEPY 736
Db 661 VGVGCGGSSLSLALAMDQVEGHVATKGSVAVYPOQAMTQNDLSRENILLPGCOLBEPY 720
Qy 737 YRSVYQACALPDLIELPSGDRTEIGEKGVNLSGGQKORVSLARAVYNSADIYLFDDPLS 796
Db 721 YRSVYQACALPDLIELPSGDRTEIGE----- 747
Qy 797 AVDAHVGRIPEENVIGPKMLKNKRIIIVTHSMSTYLPQVDVITIVMSGKISMSGYQELL 856
Db 748 -----KTRIIIVTHSMSTYLPQVDVITIVMSGKISMSGYQELL 784
Qy 857 ARDGAFAEFLRTYATSTEOBODAEENGVTGVSPGKEAKOMENGMVITISAGKQIOLROJSS 916
Db 785 ARDGAFAEFLRTYATSTEOBODAEENGVTGVSPGKEAKOMENGMVITISAGKQIOLROJSS 844
Qy 917 SSSYSGDLSRRHNSYAELOKAEKKEETWKLEAKAQTQYKLSVYMDYMKALGLFISF 976
Db 845 SSSYSGDLSRRHNSYAELOKAEKKEETWKLEAKAQTQYKLSVYMDYMKALGLFISF 904
Qy 977 LSLIFLMCNHVSALASNYWLSLMTDPIVNGTQEHKVLVLSYGAIGISQGIAGVGYGMA 1036
Db 905 LSLIFLMCNHVSALASNYWLSLMTDPIVNGTQEHKVLVLSYGAIGISQGIAGVGYGMA 964
Qy 1037 VSIIGIILASRCLAHVLLHSLILSIPMSFPERTSGULVNRFSKELDTTVDSMIPEVTKMFG 1096
Db 965 VSIIGIILASRCLAHVLLHSLILSIPMSFPERTSGULVNRFSKELDTTVDSMIPEVTKMFG 1024
Qy 1097 SLFNVIGACIVILLATPIAAIILPPLGLIYFPQRFYVVASQOLKRLSVSRSPYASHPN 1156
Db 1025 SLFNVIGACIVILLATPIAAIILPPLGLIYFPQRFYVVASQOLKRLSVSRSPYASHPN 1084
Qy 1157 ETLGLVSVIRAEEOERFIHQSDLVADENOKAYPSIVANRWLAVALRLECVCNACIVLFAAL 1216
Db 1085 ETLGLVSVIRAEEOERFIHQSDLVADENOKAYPSIVANRWLAVALRLECVCNACIVLFAAL 1144
Qy 1217 FAVISRHSISAGLVGLSVYSYSLQVTTYLANVLVYRMSSEMETNIVAVBRLEKYSETEKEAPW 1276
Db 1145 FAVISRHSISAGLVGLSVYSYSLQVTTYLANVLVYRMSSEMETNIVAVBRLEKYSETEKEAPW 1204
Qy 1277 QIQETAPSSWPQVQVREFRNCLARYREDLDVYLHINVTINGEKGVGVIGRTGAGKSL 1336
Db 1205 QIQETAPSSWPQVQVREFRNCLARYREDLDVYLHINVTINGEKGVGVIGRTGAGKSL 1264
Qy 1337 TLGLRFINESAGEIILIDGINIAKIGLHDLRKITIIIPQDPVLFSGSLRMLNDPFSQYSD 1396
Db 1265 TLGLRFINESAGEIILIDGINIAKIGLHDLRKITIIIPQDPVLFSGSLRMLNDPFSQYSD 1324
Qy 1397 BEVWVTSLELAHKDQVYSALPDKLDHCECAGGENLSVGQOLVCLARALLRKTKIIVLDBA 1456
Db 1325 BEVWVTSLELAHKDQVYSALPDKLDHCECAGGENLSVGQOLVCLARALLRKTKIIVLDBA 1384

Qy 1457 TAAVDETDLLIGSTIRTOEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILQ 1516
Db 1385 TAAVDETDLLIGSTIRTOEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILQ 1444
Qy 1517 ORGLFYGMADQAGLV 1531
Db 1445 ORGLFYGMADQAGLV 1459
RESULT 6
Q9U0A0 HUMAN
ID Q9U0A0 HUMAN PRELIMINARY; PRT; 1456 AA.
AC Q9U0A0;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TRENBLrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.F., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF022853; AAB83981.1; -; Genomic DNA.
DR EMBL; AF022824; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022826; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022827; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022829; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022831; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022840; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022838; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022852; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022851; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022849; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022848; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022828; AAB83981.1; JOINED; Genomic DNA.
DR HSSP; P08716; IMTO.
DR Ensembl; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transpc.
DR InterPro; IPR003439; ABC_tranep_like.

DR InterPro: IPR005292; MRP assoc.
 DR InterPro: IPR000719; Prot. kinase.
 DR Pfam: PF00664; ABC_membrane_2.
 DR Pfam: PF00005; ABC_tran_2.
 DR ProDom: PD00006; ABC_transporter_2.
 DR SMART: SM00382; AAA_2.
 DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE: PS0929; ABC_TMIF_2; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE: PS08993; ABC_TRANSPORTER_2; 2.
 DR PROSITE: PS0107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR ATP-binding; Nucleotide-binding.
 FT NON TER 1
 SQ SEQUENCE 1456 AA; 163233 MW; 8DB8AB2B0481F2 CRC64;

Query Match 74.9%; Score 7419.5; DB 2; Length 1456;
 Best Local Similarity 96.1%; Pred. No. 0;
 Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;

QY 17 DMNTNTNSNDPFTKQNTVTVVWPCFYLMACFPFYLYSRHDSGTYQMTPLNKTKTA 76
 DB 1 DMNTNTNSNDPFTKQNTVTVVWPCFYLMACFPFYLYSRHDSGTYQMTPLNKTKTA 60

QY 77 LGFLMTVCMADLFYSEWERSRGIFLAPVPLVSPFLIGITTLATFLIQLERRKGVSSG 136
 DB 61 LGFLMTVCMADLFYSEWERSRGIFLAPVPLVSPFLIGITTLATFLIQLERRKGVSSG 120

QY 137 IMLTFWLVAVCALAIRSKIMTALKEDAQVDLFRDITFYVYFSLLIQVLSCFSDRSP 196
 DB 121 IMLTFWLVAVCALAIRSKIMTALKEDAQVDLFRDITFYVYFSLLIQVLSCFSDRSP 180

QY 197 LFSTIHDPNCPSSASFLSRITFWMTGLIVRGYQPLEGSLMSLNKEDTSEQVVPV 256
 DB 181 LFSTIHDPNCPSSASFLSRITFWMTGLIVRGYQPLEGSLMSLNKEDTSEQVVPV 240

QY 257 LVKMKKECAKTRKOPKVVVYSSKDPAPKESKVDANEVEALIVKSPKEMPSLPEKV 316
 DB 241 LVKMKKECAKTRKOPKVVVYSSKDPAPKESKVDANEVEALIVKSPKEMPSLPEKV 300

QY 317 LYKTFGYFLMSFPFKAIDHLMFSGPQILKLIKPVNDTAPDMOGFYTVLLFVTACL 376
 DB 301 LYKTFGYFLMSFPFKAIDHLMFSGPQILKLIKPVNDTAPDMOGFYTVLLFVTACL 360

QY 377 QTVLHGYFHCFSGGRITKAVIGAYYRKALVITNSRKSTYGEIYNLMSVDAQRPMD 436
 DB 361 QTVLHGYFHCFSGGRITKAVIGAYYRKALVITNSRKSTYGEIYNLMSVDAQRPMD 420

QY 437 LATYINNIWSAPLOVIALYLMNLGSPVLAVAVWVAVVAVVAVVAVVAVVAVVAVVAVV 496
 DB 421 LATYINNIWSAPLOVIALYLMNLGSPVLAVAVVAVVAVVAVVAVVAVVAVVAVVAVV 480

QY 497 SKDNRIKLMELINGIKVILKLYAMELAFKDKVLAIROELKVLKKSAYLSAVGTFTWCT 556
 DB 481 SKDNRIKLMELINGIKVILKLYAMELAFKDKVLAIROELKVLKKSAYLSAVGTFTWCT 540

QY 557 PRLVALCTPAAVYITIDENNIIDAOATAPVSLAFNITLAPPLNITLPMVSSITQASVSLKRL 616
 DB 541 PRLVALCTPAAVYITIDENNIIDAOATAPVSLAFNITLAPPLNITLPMVSSITQASVSLKRL 600

QY 617 RIFLSHELEPDSIERRPVKDGGTNSITVNNATFTVARSPTLNGITGSPGALVAV 676
 DB 601 RIFLSHELEPDSIERRPVKDGGTNSITVNNATFTVARSPTLNGITGSPGALVAV 660

QY 677 VGVGVCKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOAWIIONDSLRENILLFGQLEBPY 736
 DB 661 VGVGVCKSSLLSALLAEMDKVEGHVAI----- 688

QY 737 YRSVYQACALLPDELILPSGDRTEIGEGVNLSSGQKORVSLABAVYSNADIYLPDPLS 796
 DB 689 -----KGVNLSSGQKORVSLABAVYSNADIYLPDPLS 721

QY 797 AVDAHVGHKIFENYIGPKMLKNKTRILVTHSMGYLPQVDVYIYWSGGKISEMSYQEL 856

DB 722 AVDAHVGHKIFENYIGPKMLKNKTRILVTHSMGYLPQVDVYIYWSGGKISEMSYQEL 781

QY 857 ARDGAFAEPLRTYASTQEDAEENGYTGVSQPKKAKQKENGMLVYDSACKQIQORQLS 916

DB 782 ARDGAFAEPLRTYASTQEDAEENGYTGVSQPKKAKQKENGMLVYDSACKQIQORQLS 841

QY 917 SSSYSGISRRHNSTAELOKAEAKKETWKLMEADKQOTQOVKLSVYWDYKAIQGLFISF 976

DB 842 SSSYSGISRRHNSTAELOKAEAKKETWKLMEADKQOTQOVKLSVYWDYKAIQGLFISF 901

QY 977 LSIFLPMCNHVSALASNYWMLSLWTDPIVNGQDHTKVRLSVYGALGISQIAVFGYSMA 1036

DB 902 LSIFLPMCNHVSALASNYWMLSLWTDPIVNGQDHTKVRLSVYGALGISQIAVFGYSMA 961

QY 1037 VSIIGILASRCLHYDLHSLRSPMSFERTPSGNLVNRSKELDYDSMTPEYIKNFMG 1096

DB 962 VSIIGILASRCLHYDLHSLRSPMSFERTPSGNLVNRSKELDYDSMTPEYIKNFMG 1021

QY 1097 SLFNVIGACIVILLATPIAAIIPPLGLIYFVQRFVVASRQLKLESYSRSPVSGHN 1156

DB 1022 SLFNVIGACIVILLATPIAAIIPPLGLIYFVQRFVVASRQLKLESYSRSPVSGHN 1081

QY 1157 ETLGVSIVIRAFEOERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCTVLPAL 1216

DB 1082 ETLGVSIVIRAFEOERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCTVLPAL 1141

QY 1217 PAVISRHSISAGLVGSLVSYSLQVTTYLNLVWSSSEMETNIVAVERLKEYSETEKAPW 1276

DB 1142 PAVISRHSISAGLVGSLVSYSLQVTTYLNLVWSSSEMETNIVAVERLKEYSETEKAPW 1201

QY 1277 QIOETAPSSMPQVGRFEPNVCYRVEDLDPVLRHINVTINGEKYGIYRTAGAGSSL 1336

DB 1202 QIOETAPSSMPQVGRFEPNVCYRVEDLDPVLRHINVTINGEKYGIYRTAGAGSSL 1261

QY 1337 TLGFLRINESAEGEIIIDGINIAKIGHDLRFKTIIPQDPVLFSGSLRMLDPPSOYSD 1396

DB 1262 TLGFLRINESAEGEIIIDGINIAKIGHDLRFKTIIPQDPVLFSGSLRMLDPPSOYSD 1321

QY 1397 EEVWTSLELAHLKDFVSALPDKLDHBCAEGENLSVGOROLVCLARALLRKTILVDEA 1456

DB 1322 EEVWTSLELAHLKDFVSALPDKLDHBCAEGENLSVGOROLVCLARALLRKTILVDEA 1381

QY 1457 TAAVDLETDLLIOSTRTOFEDCTVLTARLNTIMTYRVIVYVDKELQIYEGAPSLQ 1516

DB 1382 TAAVDLETDLLIOSTRTOFEDCTVLTARLNTIMTYRVIVYVDKELQIYEGAPSLQ 1441

QY 1517 QRGFLYSMAKDAGLV 1531

DB 1442 QRGFLYSMAKDAGLV 1456

RESULT 7
 Q6UR05 CANPA
 ID Q6UR05 CANPA PRELIMINARY; PRT; 1531 AA.
 AC Q6UR05;
 DT 05-JUL-2004 (Tremblere1. 27, Created)
 DT 05-JUL-2004 (Tremblere1. 27, Last sequence update)
 DT 05-JUL-2004 (Tremblere1. 27, Last annotation update)
 DE Multidrug resistance-associated protein 1.
 GN Name=MRP1;
 OS Canis familiaris (dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 NC NCB1_TaxID=9615;
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=22404678; PubMed=12516967;
 RA Ma L., Pratt S.B., Cao J., Danczig A.H., Moore R.E., Slapak C.A.;
 RT "Identification and characterization of the canine multidrug
 resistance-associated protein.";
 RL Mol. Cancer Ther. 1:1335-1342(2002).
 [2]

RP NUCLEOTIDE SEQUENCE.
Ma L., Pratt S.E., Cao J., Danczig A.H., Moore R.E., Slapak C.A.;
Submitted (Aug-2003) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AY633728; AAC33148.1; -; mRNA.
DR Ensembl: ENSCAFG00000018206; Canis familiaris.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042656; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transporter; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR011527; ABC_membrane_1.
DR InterPro: IPR011440; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR005292; MRP_assoc.
DR InterPro: IPR000719; Proc_kinase.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS0929; ABC_TMTP_2; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS08993; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS0107; PROTEIN_KINASE_ATP; UNKNOWN 1.
KW ATP-binding; Nucleotide-binding; Repeat; Transport.
SQ SEQUENCE 1531 AA; 171792 MW; 746361A71C6158BD CRC64;

Query Match 73.4%; Score 7272; DB 2; Length 1531;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;

QY 1 MALGFGCAGDSDDPLMDNMVMTNTSNPDTFCQONTVAVWPCFYLMACFPPEFLYSRHH 60
DB 1 MALGFGCAGDSDDPLMDNMVMTNTSNPDTFCQONTVAVWPCFYLMACFPPEFLYSRHH 60
QY 61 DRGYIOWTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLTGITTLA 120
DB 61 DRGYIOWTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLTGITTLA 120
QY 121 TFLIQLEERRKGVSSGIMLTFMLVALVCALALIRSKIMTALKEDAOVDLFRDITFYVYS 180
DB 121 TFLIQLEERRKGVSSGIMLTFMLVALVCALALIRSKIMTALKEDAOVDLFRDITFYVYS 180
QY 181 LLLIQVLVSCSDRSPLFSETHIDNPCESSASFLSITFWMTGLVRCGRQLESDD 240
DB 181 LLLIQVLVSCSDRSPLFSETHIDNPCESSASFLSITFWMTGLVRCGRQLESDD 240
QY 241 LMSLNKEDTSQOVVAVLVKNMKKECAKTRKQPKVAVVSSKDPAPKRESSKYDANEVEAL 300
DB 241 LMSLNKEDTSQOVVAVLVKNMKKECAKTRKQPKVAVVSSKDPAPKRESSKYDANEVEAL 300
QY 301 IVKSPQKEMWNSLFVLYKTFGPYFLMSFFKAIHDLMMFSGPQILKLIFVNDYKAPD 360
DB 301 IVKSPQKEMWNSLFVLYKTFGPYFLMSFFKAIHDLMMFSGPQILKLIFVNDYKAPD 360
QY 361 WOGFYTYVLFEVTLQLOTVLHQVFHICFVSGMRKTAIVIGAVYKAVITNNSAKSSSTV 420
DB 361 WOGFYTYVLFEVTLQLOTVLHQVFHICFVSGMRKTAIVIGAVYKAVITNNSAKSSSTV 420
QY 421 GEIYVIMSVDAQRFMDLATYINMISAPLOVIALYLMLNIGSPVLAVGAVMVLVNVN 480
DB 421 GEIYVIMSVDAQRFMDLATYINMISAPLOVIALYLMLNIGSPVLAVGAVMVLVNVN 480
QY 481 AVAMKTKTYQVAHKKSKDNRIKLMEILNGIKVLKJYAMELAFKQKYLAIROBELKYLK 540
DB 481 AVAMKTKTYQVAHKKSKDNRIKLMEILNGIKVLKJYAMELAFKQKYLAIROBELKYLK 540
QY 541 KSAIYLSANGTFTWCTPFLVALCTFAVYVITDENNILLDAQTFVSLAFNIRFLANTILP 600
DB 541 KSAIYLSANGTFTWCTPFLVALCTFAVYVITDENNILLDAQTFVSLAFNIRFLANTILP 600

QY 601 MVISSIVQASVSLKRLRILFISHSELEPDSIERRPVKDGGGNTSITVENAFTTWARSDPPT 660
DB 601 MVISSIVQASVSLKRLRILFISHSELEPDSIERRPVKDGGGNTSITVENAFTTWARSDPPT 660
QY 661 LNGTHSPREGALVAVNGVQCCGSSLSLSALLAMDVVEGVALKSGVAVVPOQAMIOND 720
DB 661 LNGTHSPREGALVAVNGVQCCGSSLSLSALLAMDVVEGVALKSGVAVVPOQAMIOND 720
QY 721 SLRNLIFGCOLLEPPYRSVYIQAALLPDEILPSCGRTEIGEGKVLNSGGOKORVSLAR 780
DB 721 SLRNLIFGCOLLEPPYRSVYIQAALLPDEILPSCGRTEIGEGKVLNSGGOKORVSLAR 780
QY 781 AVVSNADIIYLPDDLGAVDHVGKHFENYIGPKMLKNKTRILIVTSMSTLPQVDVYIV 840
DB 781 AVVCDSDIYLPDDLGAVDHVGKHFENYIGPKMLKNKTRILIVTSMSTLPQVDVYIV 840
QY 841 MSGKISEMGSYQELLARDDGAFAEFLTYASTBOEDABENGVAVSGPGKEAKOMENG 900
DB 841 MTGKISEMGSYQELLARDDGAFAEFLTYASTBOEDABENGVAVSGPGKEAKOMENG 900
QY 901 LVYDSAGKOLQRQSSSSSYSGDISRHHNSTAELOKAFAKKEETWKLMEADKATQGVYKL 960
DB 901 LVYDSAGKOLQRQSSSSSYSGDISRHHNSTAELOKAFAKKEETWKLMEADKATQGVYKL 960
QY 961 SVYWDYMKALGLFISLIFLEMCNHVASALASNTYLSLWTDPIVNGTOEHTKYRLSYVG 1020
DB 961 SVYWDYMKALGLFISLIFLEMCNHVASALASNTYLSLWTDPIVNGTOEHTKYRLSYVG 1020
QY 1021 ALGISQGIYAVRGYMAVSIIGITLASRCHLDLHSLIRSPMSFERPSPGULVNRFSKEL 1080
DB 1021 ALGISQGIYAVRGYMAVSIIGITLASRCHLDLHSLIRSPMSFERPSPGULVNRFSKEL 1080
QY 1081 DTVDMSMPEVIMKMGSLFENVIGACIVLATAPIIAIIPPLGLIYFEVQRFYVASSRQL 1140
DB 1081 DTVDMSMPEVIMKMGSLFENVIGACIVLATAPIIAIIPPLGLIYFEVQRFYVASSRQL 1140
QY 1141 KRLESVSRSPVYSHFNETLGVSYIYRAFEQERFIHOSDLKVDENOKAYYPSIVANRWLA 1200
DB 1141 KRLESVSRSPVYSHFNETLGVSYIYRAFEQERFIHOSDLKVDENOKAYYPSIVANRWLA 1200
QY 1201 VRLCEVNGCIVLPALPAVISRHSLSAGVLGSYSIQVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNGCIVLPALPAVISRHSLSAGVLGSYSIQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEVSETEKAPWQIOETAPSPSPQVGRVFRNYCLTRYREDLDFVLHINVTINGG 1320
DB 1261 VERLKEVSETEKAPWQIOETAPSPSPQVGRVFRNYCLTRYREDLDFVLHINVTINGG 1320
QY 1321 EKVGIYGRTGAKSLLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVGIYGRTGAKSLLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
QY 1381 SGSLRPMNIDPFSQVSDERWVSLBLAHKDVSAALPQDLBHECAEGENLSVGGORQVYL 1440
DB 1381 SGSLRPMNIDPFSQVSDERWVSLBLAHKDVSAALPQDLBHECAEGENLSVGGORQVYL 1440
QY 1441 ARALLRKTIVLBEATAVDELTDLDLQSTIRFQFEDCTVLTJAHRLANTIMDYTRYVYL 1500
DB 1441 ARALLRKTIVLBEATAVDELTDLDLQSTIRFQFEDCTVLTJAHRLANTIMDYTRYVYL 1500
QY 1501 DKGEIOEYGAPSDILLQGRGLFYSSAKDAGLY 1531
DB 1501 DKGEIOEYGAPSDILLQGRGLFYSSAKDAGLY 1531

RESULT 8
059G19 HUMAN
ID 059G19 HUMAN PRELIMINARY; PRT; 1439 AA.
AC 059G19
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)

DE ATP-binding cassette, sub-family C, member 1 isoform 1 variant
(Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
ON NCBI_TaxID=9606;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain.
RA TOOKI Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.,
RT "None Title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SMILLIARITY: Belongs to the ABC transporter family.
DR EMBL; AB209120; BAD92357.1; -; mRNA.
DR Ensemble; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR004339; ABC_TM_transp.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding.
FT NON_TER 1
SQ SEQUENCE 1439 AA; 160439 MW; 090D881ADBFB3077 CRC64;
Query Match 73.1%; Score 7242; DB 2; Length 1439;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1428; Conservative 0; Mismatches 1; Indels 10; Gaps 1;
QY 103 APVFLVSPPTLLIGITTLATFLIQLERRKGVSQSIMLTFMVLVALVCALAILRSKIMTALK 162
DB 1 APVFLVSPPTLLIGITMLATFLIQLERRKGVSQSIMLTFMVLVALVCALAILRSKIMTALK 60
QY 163 EDNOVDLFRDITFVYVSSLLIQVLVSCFSDRSPLFSETHIDPRCPRESSASPLSRITFW 222
DB 61 EDNOVDLFRDITFVYVSSLLIQVLVSCFSDRSPLFSETHIDPRCPRESSASPLSRITFW 120
QY 223 WITGLIYRGYROPLEGSDLSINKEDTSEQVPLVYKMKKECAKTRKOPKYVYSSKDP 282
DB 121 WITGLIYRGYROPLEGSDLSINKEDTSEQVPLVYKMKKECAKTRKOPKYVYSSKDP 180
QY 283 AQPRESSKVDANEVEALIVKSPQKWNPSLFFVLYKTFGPFYLFMSFFFAIHDLMVFSG 342
DB 181 AQPRESSKVDANEVEALIVKSPQKWNPSLFFVLYKTFGPFYLFMSFFFAIHDLMVFSG 240
QY 343 POLIKLLIKFVNPKADNGCYFTYVLLFTYACLQTLVLAQYFHTICVSGMRITAVIGA 402
DB 241 POLIKLLIKFVNPKADNGCYFTYVLLFTYACLQTLVLAQYFHTICVSGMRITAVIGA 300
QY 403 VYRKALVITNSARKSSVTGEIVNLMSYDAQRFMDLATYINMISAPLOVILLATYILMLNL 462
DB 301 VYRKALVITNSARKSSVTGEIVNLMSYDAQRFMDLATYINMISAPLOVILLATYILMLNL 360
QY 463 GPSTLAVAVMTLVNPNVAMAKTKTYOVAHMSKNRIKLNELINGIKVLKYAWEL 522
DB 361 GPSTLAVAVMTLVNPNVAMAKTKTYOVAHMSKNRIKLNELINGIKVLKYAWEL 420

QY 523 AFKDKVLAIOBELKVLKKSAYVLSAVGTFTWCTPFLVALCTFAVYVYIDENNILDAQTA 582
DB 421 AFKDKVLAIOBELKVLKKSAYVLSAVGTFTWCTPFLVALCTFAVYVYIDENNILDAQTA 480
QY 583 FVSLALFNILRFPNLILPMVLISSIVQASVSLKRLIFLSHELEPDSIERRPVQGGGTN 642
DB 481 FVSLALFNILRFPNLILPMVLISSIVQASVSLKRLIFLSHELEPDSIERRPVQGGGTN 540
QY 643 SITVRNATFTWASDPTLNGITPSPREGALVAVVGVGGKSLSLALLAENDKVEGHV 702
DB 541 SITVRNATFTWASDPTLNGITPSPREGALVAVVGVGGKSLSLALLAENDKVEGHV 600
QY 703 AIKGSVAVYVQOAMQWDSIRENILEGCOLPEEPPYRSVIOACALLPDLTLPSCDREIG 762
DB 601 AIKGSVAVYVQOAMQWDSIRENILEGCOLPEEPPYRSVIOACALLPDLTLPSCDREIG 660
QY 763 EKGVLSSGQKORVSLARAVYNSADITLFPDPLSAVDAAHVKHIFENVIGPKMLKNKTR 822
DB 661 EKGVLSSGQKORVSLARAVYNSADITLFPDPLSAVDAAHVKHIFENVIGPKMLKNKTR 720
QY 823 ILVTHSMSTYLPQVDVITVMSGGKISEMGSYQELIARDGAPAEFLRTYASTEOBDAEEN- 881
DB 721 ILVTHSMSTYLPQVDVITVMSGGKISEMGSYQELIARDGAPAEFLRTYASTEOBDAEEN- 780
QY 882 -----GVTVSGPKGKAKOMENGMLVTDASAGKQLORQLSSSSSGSDISRHNSTA 932
DB 781 STWDEEEAAGVTVSGGKAKOMENGMLVTDASAGKQLORQLSSSSSGSDISRHNSTA 840
QY 933 ELQKAEKKEETKWLMEADRAQTGVQLSYWMDYMKAGHIFISFLIFLMCHVSLAAS 992
DB 841 ELQKAEKKEETKWLMEADRAQTGVQLSYWMDYMKAGHIFISFLIFLMCHVSLAAS 900
QY 993 NYWLSMTDDPIYNGTOEHTKVLRYVYALGISQGIARVFGYSMAVSTIGILASRCHVDL 1052
DB 901 NYWLSMTDDPIYNGTOEHTKVLRYVYALGISQGIARVFGYSMAVSTIGILASRCHVDL 960
QY 1053 LHSILSPMSFERTSGNLYNRFSEKELDVDSIMEBVIKMFNGSLFNVIAGCIVILLAT 1112
DB 961 LHSILSPMSFERTSGNLYNRFSEKELDVDSIMEBVIKMFNGSLFNVIAGCIVILLAT 1020
QY 1113 PIAAIIIPPLGLIYFFQRFYVASSRQLKRLSESVRSPPVSHFNETLLAGSVIRAFPEOE 1172
DB 1021 PIAAIIIPPLGLIYFFQRFYVASSRQLKRLSESVRSPPVSHFNETLLAGSVIRAFPEOE 1080
QY 1173 RFIHQSDLKVDENQKAYPSIVANRWLAVALCEVNCIVLFAALPAVISRHSLSAGLVGL 1232
DB 1081 RFIHQSDLKVDENQKAYPSIVANRWLAVALCEVNCIVLFAALPAVISRHSLSAGLVGL 1140
QY 1233 SVSISLQVTTYLWMLVRMSSEMETNIYAVERLKEYSETEKEAPWQIOETAPSSWPQVR 1292
DB 1141 SVSISLQVTTYLWMLVRMSSEMETNIYAVERLKEYSETEKEAPWQIOETAPSSWPQVR 1200
QY 1293 VEFRRNYCLARYEDLDVLRHINVTINGEKVGIYGRTAGKSLTGLFRINSAGEIIL 1352
DB 1201 VEFRRNYCLARYEDLDVLRHINVTINGEKVGIYGRTAGKSLTGLFRINSAGEIIL 1260
QY 1353 IDGINIAKIGLHDLRFKRTIIPQDPVLPFGSLRMLNDPESQYSDEEVTWLSLELAHLKDFV 1412
DB 1261 IDGINIAKIGLHDLRFKRTIIPQDPVLPFGSLRMLNDPESQYSDEEVTWLSLELAHLKDFV 1320
QY 1413 SALPDKLHDEHCAAGENLSVGRQVLCARALLRKTKIIVLDATAVVDLETDDLQSTI 1472
DB 1321 SALPDKLHDEHCAAGENLSVGRQVLCARALLRKTKIIVLDATAVVDLETDDLQSTI 1380
QY 1473 RTQFEDCTVLTIAHRLNTIMDYTRVIVLXKGEIOEYGAESPDLIQOGLFYSMAKDAGLV 1531
DB 1381 RTQFEDCTVLTIAHRLNTIMDYTRVIVLXKGEIOEYGAESPDLIQOGLFYSMAKDAGLV 1439
RESULT 9
Q8HXQ5_BOVIN
ID Q8HXQ5_BOVIN PRELIMINARY; PRT; 1530 AA.
AC Q8HXQ5;

DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Multidrug resistance protein 1.
GN Name=MRP1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;
RA Taguchi T., Saeki K., Komano T.;
RT "Functional analysis of MRP1 cloned from bovine."
RL FEBS Lett. 521:211-213(2002).
CC -1 - SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB082124; BAC15550.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR GO; GO:0016021; C:Integral to membrane, IEA.
DR GO; GO:0005524; F:ATP binding, IEA.
DR GO; GO:0016887; F:ATPase activity, IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding, IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001440; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF_2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR ATP-binding; Nucleotide-binding; Repeat.
KW SEQUENCE 1530 AA; 171666 MW; AAAF92ED7832703 CRC64;
SQ
Query Match 73.0%; Score 7230.5; DB 2; Length 1530;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;
QY 1 MALRGFCAGDSDDPLMDNMVNTWNTSNPDFTKCFQNTVAVWVPCFYLMACFPFFLYLSRH 60
DB 1 MALNDFCGVDSDDLFWEMVNTWNTSNPDFTKCFQNTVAVWVPCSYLMWCFPPFFLYLSRH 60
QY 61 DRGYIQMTPLNKTATGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGITTLLA 120
DB 61 DRGYIQMTPLNKTATGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGITTLLA 120
QY 121 TFLIQLERKGVQSSGIMLTFWLVALVCALALRSKIMTALKEDQVDLFRDITRYVFS 180
DB 121 TFLIQLERKGVQSSGIMLTFWLVALVCALALRSKIMTALKEDQVDLFRDITRYVFS 180
QY 121 TFLIQLERKGVQSSGIMLTFWLVALVCALALRSKIMTALKEDQVDLFRDITRYVFS 180
DB 121 TFLIQLERKGVQSSGIMLTFWLVALVCALALRSKIMTALKEDQVDLFRDITRYVFS 180
QY 181 LLLIQLVSCFSDRSPLFSETLHDPNCPESASFLSITTFWMTGLIVRGROPLEGSD 240
DB 181 LLLIQLVSCFSDRSPLFSETLHDPNCPESASFLSITTFWMTGLIVRGROPLEGSD 240
QY 241 LMSLNKEDTSQVAVLVKNNKKECAKTRKQPVKVVYSSKDPAPQKSSKVDANEVVAL 300
DB 241 LMSLNKEDTSQVAVLVKNNKKECAKTRKQPVKVVYSSKDPAPQKSSKVDANEVVAL 300
QY 301 IVKSPQKEMNPDLFLVLYKTGPPYFLMSFFPKAHHDLMMFSGPOLIKLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPDLFLVLYKTGPPYFLMSFFPKAHHDLMMFSGPOLIKLIFVNDTKAPD 360
QY 361 MGGFYFYLFLPYTACLOTLVLMHQVHICVSGMRKTAVIGAVYKALVITMSARKSSSTV 420
DB 361 MGGFYFYLFLPYTACLOTLVLMHQVHICVSGMRKTAVIGAVYKALVITMSARKSSSTV 420

QY 421 GEIYNLMSVDAQRMDLATTYNNMTWSPAPLOYITLALYLLMLNLGSSVLAGVAVNMLAMPVN 480
DB 421 GEIYNLMSVDAQRMDLATTYNNMTWSPAPLOYITLALYLLMLNLGSSVLAGVAVNMLAMPVN 480
QY 481 AVMAKTKTYOVAMKSKDNRIKLMNEILINGIKYIKIYAMELAKDKYVLAIRQELKYLK 540
DB 481 AVMAKTKTYOVAMKSKDNRIKLMNEILINGIKYIKIYAMELAKDKYVLAIRQELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIDAQTAFFVSLAFNIREPPLNLP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIDAQTAFFVSLAFNIREPPLNLP 600
QY 601 MVISIYQASVSLKRLKILFISHERLEPDSIFERRPVKQCGGNTSTVNNATFTWASDPPT 660
DB 601 MVISIYQASVSLKRLKILFISHERLEPDSIFERRPVKQCGGNTSTVNNATFTWASDPPT 660
QY 661 LMGITFSPGALVAVVGVGCGKSSLLSLALAMDYVEGHVALKGSVAVYPOQAMIOND 720
DB 661 LMGITFSPGALVAVVGVGCGKSSLLSLALAMDYVEGHVALKGSVAVYPOQAMIOND 720
QY 721 SLRENILFGCOLBEPYRSVYIQAALLPDEILPBGDRTEIGEGVNLGGQKQSVLSAR 780
DB 721 SLRENILFGCOLBEPYRSVYIQAALLPDEILPBGDRTEIGEGVNLGGQKQSVLSAR 780
QY 781 AVYSNADITYLPDDLAVDAVAGHIFBNYIGPKGMLKNTKRLIVYTHSMSTLPQVDYIY 840
DB 781 AVYCDSDVYLLDDPLSAVDVAGHIFBNYIGPKGMLKNTKRLIVYTHAISYLPQVDYIY 840
QY 841 MSGGKISPMGSYQELRLARDGAFAEFLRTYASTBEOBDAEENGVYSGPGKEAOMENGM 900
DB 841 MSGGKISPMGSYQELRLARDGAFAEFLRTYASTBEOBDAEENGVYSGPGKEAOMENGM 900
QY 901 LVYDSAGKOLQRLSSSSSYSGDISRHNSYAELOKA BAKKEETWKLMEADKAQTQVYL 960
DB 901 LVYDSAGKOLQRLSSSSSYSGDISRHNSYAELOKA BAKKEETWKLMEADKAQTQVYL 960
QY 961 SVYWDYKAIQLFISFISIFLPMCNHVSALASNYWLSMTDDPLVNGTOEHTKRLSYG 1020
DB 961 SVYWDYKAIQLFISFISIFLPMCNHVSALASNYWLSMTDDPLVNGTOEHTKRLSYG 1020
QY 1021 ALGISOGIAVGYSMAYSIGIILASRCILHDLHSIIRSPSPFERPSSGLVNRFSKEL 1080
DB 1021 ALGISOGIAVGYSMAYSIGIILASRCILHDLHSIIRSPSPFERPSSGLVNRFSKEL 1080
QY 1080 DTVSMIPEVIMKMGSLFNVIAGACIYLLATPIAIIIPPLGIYFFVQGFYVASSRQL 1140
DB 1080 DTVSMIPEVIMKMGSLFNVIAGACIYLLATPIAIIIPPLGIYFFVQGFYVASSRQL 1140
QY 1141 KRLSVSRSPVYSHFNETLIGVSVITRAFEQGERFIHOSDLKVDENQKAYPSIYANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLIGVSVITRAFEQGERFIHOSDLKVDENQKAYPSIYANRWLA 1200
QY 1201 VRLCEVNGCIYLPALPAVISRHSLSAGIVGLSVSYSLQVTTYNNMLVRMSEMETNIVA 1260
DB 1201 VRLCEVNGCIYLPALPAVISRHSLSAGIVGLSVSYSLQVTTYNNMLVRMSEMETNIVA 1260
QY 1261 VERLKEVSETEKAPMIOETAPPSWPOVGRVFFRNYCLATREDDLFLVLAHINVTING 1320
DB 1261 VERLKEVSETEKAPMIOETAPPSWPOVGRVFFRNYCLATREDDLFLVLAHINVTING 1320
QY 1320 VERLKEVSETEKAPMIOETAPPSWPOVGRVFFRNYCLATREDDLFLVLAHINVTING 1320
DB 1320 VERLKEVSETEKAPMIOETAPPSWPOVGRVFFRNYCLATREDDLFLVLAHINVTING 1320
QY 1381 SGSLRPNLDPFSQVSDERWTSLELAHKDVSALPKDLHECBGGENLSVGOQOLVCL 1440
DB 1381 SGSLRPNLDPFSQVSDERWTSLELAHKDVSALPKDLHECBGGENLSVGOQOLVCL 1440
QY 1441 ARALLRKTILVLEBATVAVDLETDDLIQSTIRFOPEDCYLTIAHRLNTIMDTTRVYL 1500
DB 1441 ARALLRKTILVLEBATVAVDLETDDLIQSTIRFOPEDCYLTIAHRLNTIMDTTRVYL 1500
QY 1501 DKGEIOEYGAPEDDLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPEDDLQORGLFYSMAKDAGLV 1531

Qy 977 LSLFLMCMNVSAALASNTWLSLMTDDPLVNGTOETKRLSLVYGALGISOGIANFPGYSMA 1036
 Db 846 LSLFLMCMNVSAALASNTWLSLMTDDPLVNGTOETKRLSLVYGALGISOGIANFPGYSMA 905
 Qy 1037 VSIIGILLASRCGLHYDLHSIIIRSPMSFEPTPSGNLVRFSKEDTVDVSMIPEVIXOMMG 1096
 Db 906 VSIIGILLASRCGLHYDLHSIIIRSPMSFEPTPSGNLVRFSKEDTVDVSMIPEVIXOMMG 965
 Qy 1097 SLFNVIGACIYLLATPIAIIIPPLGLIYFVQRFYVASSRQLKRLSVSRSPYSHFN 1156
 Db 966 SLFNVIGACIYLLATPIAIIIPPLGLIYFVQRFYVASSRQLKRLSVSRSPYSHFN 1025
 Qy 1157 ETLGLSVTRAFEEBERFIHQSDLKVDENOKAYYSIYANRWLAVRLCEVGNCLYLPAL 1216
 Db 1026 ETLGLSVTRAFEEBERFIHQSDLKVDENOKAYYSIYANRWLAVRLCEVGNCLYLPAL 1085
 Qy 1217 FAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIIVAVERLKEYSETEKAPW 1276
 Db 1086 FAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIIVAVERLKEYSETEKAPW 1145
 Qy 1277 QIOETAPPSWPQVGRVEFRNYCLYRREDLPVLRIHINVTNGEKGVIWGRTAGKSSL 1336
 Db 1146 QIOETAPPSWPQVGRVEFRNYCLYRREDLPVLRIHINVTNGEKGVIWGRTAGKSSL 1205
 Qy 1337 TLGLFRINESABGEIITIGINIAKIGLHDLRFKTIITIPQDPVLFSGSLRMULDPPSQYSD 1396
 Db 1206 TLGLFRINESABGEIITIGINIAKIGLHDLRFKTIITIPQDPVLFSGSLRMULDPPSQYSD 1265
 Qy 1397 EEWMTSLFLAHKDPVSLPDLKHECAEGGENLSVGOLOVCLARALLRKTLVLDEA 1456
 Db 1266 EEWMTSLFLAHKDPVSLPDLKHECAEGGENLSVGOLOVCLARALLRKTLVLDEA 1325
 Qy 1457 TAAVDLETDDLIQSTIRTOFEDCVLTIAHRLNTIMDYRVILDKGEIOEGXAPSDLLQ 1516
 Db 1326 TAAVDLETDDLIQSTIRTOFEDCVLTIAHRLNTIMDYRVILDKGEIOEGXAPSDLLQ 1385
 Qy 1517 QRGLEFYSMAKDAGLV 1531
 Db 1386 QRGLEFYSMAKDAGLV 1400

RESULT 11

MRP1_MOUSE STANDARD; PRT; 1528 AA.
 ID MRP1_MOUSE
 AC O35379;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).
 GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mrp;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muridae; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Muscle.
 RX MEDLINE=96251691; PubMed=8649356;
 RA Stride B.D., Valimarsion G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;
 RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter.",
 RL Mol. Pharmacol. 49:962-971(1996).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaide I., Otsu N., Saito R., Suzuki H., Yamataka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T., Baldarelli R., Hall D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schirml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chochia C., Cordani L.E., Cousins S.,
 RA Dala E., Dregan T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gasteland T., Gariboldi M., Giesi C., Godzik A., Gough J.,
 RA Giermond S., Guclincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagasima T., Nmaka K., Okido T., Pavan W.J., Petrea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sadelain A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wysshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hitozane-Kishikawa T., Kono H., Nakamura M., Sakazune N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yaenunshi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 CC -1- FUNCTION: May participate directly in the active transport of drugs into subcellular organelles or influence drug distribution indirectly (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
 CC -1- SIMILARITY: Contains 2 ABC transporter domains.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC -----
 CC EMBL; AF022908; AAC80938.1; -; mRNA.
 CC EMBL; AK029876; BAB26654.1; -; mRNA.
 CC HSSP; P08716; 1MT0.
 CC ENSEMBL; ENSMUSG00000023088; Mus musculus.
 DR MGI; MGI:102676; Abcc1.
 DR GO; GO:0005887; C:Integral to plasma membrane; IDA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR011527; ABC_membrane_1.
 DR InterPro; IPR001140; ABC_TM_transp.
 DR InterPro; IPR003439; ABC_transp_like.
 DR InterPro; IPR005292; MRP_assoc.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE; PS50929; ABC_TMRP; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
 KW ATP-binding; Glycoprotein; Nucleotide-binding; Repeat; Transmembrane; Transport.
 KM
 FT TOPO_DOM 1 33 Extracellular (By similarity).
 FT TRANSMEM 34 54 1 (By similarity).
 FT TOPO_DOM 55 74 Cytoplasmic (By similarity).
 FT TRANSMEM 75 95 2 (By similarity).
 FT TOPO_DOM 96 100 Extracellular (By similarity).
 FT TRANSMEM 101 121 3 (By similarity).
 FT TOPO_DOM 122 133 Cytoplasmic (By similarity).
 FT TRANSMEM 134 154 4 (By similarity).
 FT TOPO_DOM 155 172 Extracellular (By similarity).
 FT TRANSMEM 173 193 5 (By similarity).
 FT TOPO_DOM 194 317 Cytoplasmic (By similarity).
 FT TRANSMEM 318 338 6 (By similarity).
 FT TOPO_DOM 339 364 Extracellular (By similarity).

FT	TRANSMEM	365	385	7 (By similarity).
FT	TOPO_DOM	386	441	Cytoplasmic (By similarity).
FT	TRANSMEM	442	462	8 (By similarity).
FT	TOPO_DOM	463	465	Extracellular (By similarity).
FT	TRANSMEM	466	486	9 (By similarity).
FT	TOPO_DOM	487	548	Cytoplasmic (By similarity).
FT	TRANSMEM	549	569	10 (By similarity).
FT	TOPO_DOM	570	591	Extracellular (By similarity).
FT	TRANSMEM	592	612	11 (By similarity).
FT	TOPO_DOM	613	963	Cytoplasmic (By similarity).
FT	TRANSMEM	964	984	12 (By similarity).
FT	TOPO_DOM	985	1022	Extracellular (By similarity).
FT	TRANSMEM	1023	1043	13 (By similarity).
FT	TOPO_DOM	1044	1086	Cytoplasmic (By similarity).
FT	TRANSMEM	1087	1107	14 (By similarity).
FT	TOPO_DOM	1108	1108	Extracellular (By similarity).
FT	TRANSMEM	1109	1129	15 (By similarity).
FT	TOPO_DOM	1130	1200	Cytoplasmic (By similarity).
FT	TRANSMEM	1201	1221	16 (By similarity).
FT	TOPO_DOM	1222	1223	Extracellular (By similarity).
FT	TRANSMEM	1224	1244	17 (By similarity).
FT	TOPO_DOM	1245	1528	Cytoplasmic (By similarity).
FT	DOMAIN	326	609	ABC transmembrane type-1 1.
FT	DOMAIN	644	868	ABC transporter 1.
FT	DOMAIN	971	1253	ABC transmembrane type-1 2.
FT	DOMAIN	1290	1524	ABC transporter 2.
FT	NP_BIND	678	685	ATP 1 (potential).
FT	NP_BIND	1324	1331	ATP 2 (potential).
FT	CARBOHYD	19	19	N-linked (GlcNAc. . .) (potential).
FT	CARBOHYD	1003	1003	N-linked (GlcNAc. . .) (potential).
SO	SEQUENCE	1528 AA,	171185 MM,	68FD13667D5D1DBB CRC64,

	Query Match	Best Local Match	Similarity	80.7%;	Score 7002.5;	DB 1;	Length 1528;
	Matches 1349;	Conservative 102;	Mismatches 75;	Indels 7;	Gaps 5;		
QY	1	MALRGFCSADSDPLMDMNVMTWNTSNPDFTCCFQNTLVVWVPCFYLACPFPEFYLSRHH	60				
DB	1	MALRSFCSADSDPLMDMNVMTWNTSNPDFTCCFQNTLVVWVPCFYLACPFPEFYLSRHH	60				
QY	61	DRGIIQMTPLNKITLTAIGFLMIYCMADLFPSFMERSGIGTLAPVFLVSPLLGITTLLA	120				
DB	61	DRGIIQMTPLNKITLTAIGFLMIYCMADLFPSFMERSGIGTLAPVFLVSPLLGITTLLA	120				
QY	121	TFLLQLEBRKVOSSGIMLTFMVLAVCALILTSKIMTALKEDAQVDFPDIFRYYES	180				
DB	121	TFLLQLEBRKVOSSGIMLTFMVLAVCALILTSKIMTALKEDAQVDFPDIFRYYES	180				
QY	181	LILLIQLVLSCESDRPLFSETHDNPNCRESSASFLSRITPMWITGLIVRGYROPLEGSD	240				
DB	181	LVLVQLVLSCESDRPLFSETHDNPNCRESSASFLSRITPMWITGLIVRGYROPLEGSD	240				
QY	241	LMSLNKEDTSQVYFVLVKNMKKCAKTRKPPKYYTS-SKDPAPKRESSKYDANEEVEA	299				
DB	241	LMSLNKEDTSQVYFVLVKNMKKCAKTRKPPKYYTS-SKDPAPKRESSKYDANEEVEA	299				
QY	300	LIVSPKCKEMNPISLFKVLVYKTFEGYFLMSFEFKIHDIMMSGQIILKLLIKFINDTKAP	359				
DB	301	LIVSPKCKRBPISLFKVLVYKTFEGYFLMSFEFKIHDIMMSGQIILKLLIKFINDTKAP	360				
QY	360	DMQGYFYTVLLFVTAQLQTLVLAHQYFHLCFPSGMRIKTAVGAYRRKALLVTNSARKSST	419				
DB	361	DMQGYFYTVLLFVTAQLQTLVLAHQYFHLCFPSGMRIKTAVGAYRRKALLVTNSARKSST	420				
QY	420	VGEIVLIMSVDQAQFMDLATTYINNIWSPLOVILLATYLLMLNGLSPVLAGAVNVLAMPV	479				
DB	421	VGEIVLIMSVDQAQFMDLATTYINNIWSPLOVILLATYLLMLNGLSPVLAGAVNVLAMPV	480				
QY	480	NAVAMAKTKYQVAVHMKSKDNRIKLMLNGLIGIKVLKYAMELAFKDYVLLAIIROBELKVL	539				
DB	481	NAVAMAKTKYQVAVHMKSKDNRIKLMLNGLIGIKVLKYAMELAFKDYVLLAIIROBELKVL	540				
QY	540	KKSAVLSAVGTFTWCTPDELVALCTFAVYVTIDENNIIAQTAFAVSLAFLNIRPPLNLT	599				

Db	541	KKSAVLAAGTFTWCTPFLVALSTFAVFTVDBRNLLDACKAVSLAFENLEFPPLNLL	600
Qy	600	PMVSIIVQASVSLKRLRIFLPSHELEBPDSIEREPVDOGGSNTSYRNAFTTARSDPP	659
Db	601	PMVSIIVQASVSLKRLRIFLPSHELEBPDSIERISISGGE-NSTYKMTFTARBEPP	659
Qy	660	TLNGITTSIPFGALVAVVGQVCGCKSSLLSALLAEMDKVBGHVAIKCSAVVPOQAMION	719
Db	660	TLNGITTSIPFGALVAVVGQVCGCKSSLLSALLAEMDKVBGHVTLKCSAVVPOQAMION	719
Qy	720	DSJRENLIFGQAEPPYRSYIQACALLPDELILPSGDREIEGKGNLISGGCKORVSLA	779
Db	720	DSJRENLIFGHPLOENYKKAWEACALLPDELILPSGDREIEGKGNLSGGCKORVSLA	779
Qy	780	RAYVSNADIIYFDDPLSAVDAAHVSKHIFENYIGPKMKKNTRLTLTHSMSYLPQVYII	839
Db	780	RAYVSNADIIYFDDPLSAVDAAHVSKHIFEKVYVGMGLKKNTRLTLVHGHSYLPQVYII	839
Qy	840	VMSGKISIMSGSYOELLARDGAFAEFLRTYASTEOBDAEENGTVGVSGPCKEAKOMENG	899
Db	840	VMSGKISIMSGSYOELLARDGAFAEFLRTYANAEOQLASEDD--SYSGCKESKCPVENG	896
Qy	900	MLVYDSAGKQLOQRDSSSSSYSGDISHHNSTALQCAEAKKETTWMLEMDKAQOTQYK	959
Db	897	MLVYDVIYKGLQRRLSNSSSHSGPTSOQHSSIAELQACGA-KEETWMLMDKQAOTQOVQ	955
Qy	960	LSVYMDYMKAIIGLIFSTFSLFLPMCNVSAVSLASNYMLSLMTDD-PYNGTOEHRKVLSTV	1018
Db	956	LSVYMDYMKAIIGLFTTLFSLFLPMCNVSAVSLASNYMLSLMTDDPVNGTOANNFRLSV	1015
Qy	1019	YGALGISQGIAPVFGYSMAVSIIGLILASRCLHVDLHSLRS PMSFFERTSPGNLVNRFK	1078
Db	1016	YGALGILQGAALFQYSMAVSIIGIFASRRHLHDLVNLVSLSPMSFFERTSPGNLVNRFK	1075
Qy	1079	ELDTVDMSIPEVIMFMFNGSLFNVTGACIVILATPIAIIIIIPPLGLIYFYQRFYVASSR	1138
Db	1076	ELDTVDMSIPEVIMFMFNGSLFSYVAVIYIIILLATPIAAVYIIPPLGLVYFFQRFYVASSR	1135
Qy	1139	QLKRLSVSRSPVYSHNETLLGVSVIRAFEBEORFTHOSDKVDENOKAYPSBITVANRM	1198
Db	1136	QLKRLSVSRSPVYSHNETLLGVSVIRAFEBEORFTHOSDKVDENOKAYPSBITVANRM	1195
Qy	1199	LAVRELCVGNICIVLFAALPAVISRHSLSAGLVGLSVSYSLOVTTYLANMLVRMSSEMETNI	1258
Db	1196	LAVRELCVGNICIVLFAALPAVISRHSLSAGLVGLSVSYSLOITAYLANMLVRMSSEMETNI	1255
Qy	1259	VAVERLKEYSBTEKEAPWQIOETRAPPSMPQVGVFEFRNTCLARYREDJDPVLRHINTIN	1318
Db	1256	VAVERLKEYSBTEKEAPWQIOETRAPPSMPHSGVFEFRDYCLARYREDJDLVLKHINTIE	1315
Qy	1319	GGEVNGVIGVGTGAGKSLTGLFRINSASAGEIIIOGINAKIGLHDLRKTITIIPODPV	1378
Db	1316	GGEVNGVIGVGTGAGKSLTGLFRINSASAGEIIIDSVNIAKIGLHDLRKTITIIPODPV	1375
Qy	1379	LFSGSLTAMNIDPFSGOYDEEVMVTSLELAHKDFVSALPDKLDHECABGEGNLSVGOROLV	1438
Db	1376	LFSGSLTAMNIDPFSGOYDEEVMALBLAHKGEVSALPDKLNHECABGEGNLSVGOROLV	1435
Qy	1439	CLARALLRKTIIVLDEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVI	1498
Db	1436	CLARALLRKTIIVLDEATAVADLETDNLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVI	1495
Qy	1499	VLDKGEIOEYGAPSDILQORGLFYVMAKDXGLV 1531	
Db	1496	VLDKGEVRECGAPSELLQORGIFYMAKDXGLV 1528	
RESULT 12			
0810E4 RAT			
ID	0810E4_RAT	PRELIMINARY;	PRT; 1532 AA.
AC	0810E4;		
01-JUN-2003 (Tribble,rel. 24, Created)			
JT			

DT 01-JUN-2003 (TRENBLREL. 24, last sequence update)
DT 01-MAR-2004 (TRENBLREL. 26, last annotation update)
DE ATP-binding cassette protein C1.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
[1]
RE NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley, TISSUE=SpLeen;
RA Yabuuchi H., Takayanagi S., Ishikawa T.;
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF487549; AA085437.1; -, mRNA.
DR HSSP; P08716; IMTO.
DR Ensembl; ENSRNOG00000032748; Rattus norvegicus.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0001656; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001440; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldo/ket_red.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMR; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; UNKNOWN_1.
KM ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1532 AA; 171493 MW; 2E6939F63F5A3F68 CRC64;
Query Match 70.0%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;
QY 1 MALRGFCGADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFFLYLSRH 60
DB 1 MALSSFCSSDSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLMSCFPFLYLSRH 60
QY 61 DRGYQMTPLNKTAKALGFLMIYVQADLFYSFMRSGIPLAPVLYSPITLGLTTLLA 120
DB 61 DRGYQMTPLNKTAKALGFLMIYVQADLFYSFMRSGIPLAPVLYSPITLGLTTLLA 120
QY 121 TFLQLERKGVQSSGIMLTFWLVALVCALATLRKIMTALKEDEQVDFRDIFFYVFS 180
DB 121 TFLQLERKGVQSSGIMLTFWLVALVCALATLRKIMTALKEDEQVDFRDIFFYVFS 180
QY 121 TFLQLERKGVQSSGIMLTFWLVALVCALATLRKIMTALKEDEQVDFRDIFFYVFS 180
DB 121 TFLQLERKGVQSSGIMLTFWLVALVCALATLRKIMTALKEDEQVDFRDIFFYVFS 180
QY 181 LLLIQVLVSCFSDRSLPSETIHDPNCPRESSASFLSRITFWITGILVRYGROPLEGSD 240
DB 181 LLLIQVLVSCFSDRSLPSETIHDPNCPRESSASFLSRITFWITGILVRYGROPLEGSD 240
QY 181 LLLIQVLVSCFSDRSLPSETIHDPNCPRESSASFLSRITFWITGILVRYGROPLEGSD 240
DB 181 LLLIQVLVSCFSDRSLPSETIHDPNCPRESSASFLSRITFWITGILVRYGROPLEGSD 240
QY 241 LMSLNKETSEYQVAVLVYKMKKECAKTRKQPVKYVYS SKDPAPPKSSSKVDANEEVYA 299
DB 241 LMSLNKETSEYQVAVLVYKMKKECAKTRKQPVKYVYS SKDPAPPKSSSKVDANEEVYA 299
QY 241 LMSLNKETSEYQVAVLVYKMKKECAKTRKQPVKYVYS SKDPAPPKSSSKVDANEEVYA 299
DB 241 LMSLNKETSEYQVAVLVYKMKKECAKTRKQPVKYVYS SKDPAPPKSSSKVDANEEVYA 299
QY 300 LVIKSPKEMWNSLSEKVLVYKTFGFFLMSFFKALHDMFSGPOLIKLLIKFVNDTYAP 359
DB 300 LVIKSPKEMWNSLSEKVLVYKTFGFFLMSFFKALHDMFSGPOLIKLLIKFVNDTYAP 359
QY 301 LVIKSPKEMWNSLSEKVLVYKTFGFFLMSFFKALHDMFSGPOLIKLLIKFVNDTYAP 360
DB 301 LVIKSPKEMWNSLSEKVLVYKTFGFFLMSFFKALHDMFSGPOLIKLLIKFVNDTYAP 360
QY 360 DMQGYFTVLLFVTACLOTVLVHQYFHI CFVSGMR IKTAVIGAVYRKALVITNSARKSST 419
DB 360 DMQGYFTVLLFVTACLOTVLVHQYFHI CFVSGMR IKTAVIGAVYRKALVITNSARKSST 419

DB 361 DMQGYFTVLLFVTACLOTVLVHQYFHI CFVSGMR IKTAVIGAVYRKALVITNSARKSST 420
QY 420 VGEIVNMSVDAQRFMDLATYINMWSAPLOVILATLWNLNPGSVIAGAVVWLVV 479
DB 421 VGEIVNMSVDAQRFMDLATYINMWSAPLOVILATLWNLNPGSVIAGAVVWLVV 480
QY 480 NAWMAKTKTYOVAHMSKDNRIKLANEIJNGIKVLKUYAMELAFKDKVLAIRQELKVL 539
DB 481 NAWMAKTKTYOVAHMSKDNRIKLANEIJNGIKVLKUYAMELAFKDKVLAIRQELKVL 540
QY 540 KKSAYLSAVGTFWVCCPPLVALCTFAVYVTTDENNLIDQOTAFVSLAFNIIARFPNII 599
DB 541 KKSAYLSAVGTFWVCCPPLVALCTFAVYVTTDENNLIDQOTAFVSLAFNIIARFPNII 600
QY 600 PMVSSIVQSVSLKRLIRFLSHELEPDSIERPVQDGGTNSITVRNATFTWASDDP 659
DB 601 PMVSSIVQSVSLKRLIRFLSHELEPDSIERPVQDGGTNSITVRNATFTWASDDP 660
QY 660 TUNGITSEIGALVAVVGVGCGKSLLSALLAEMDKVGHVAKGSVAVVPOQAMION 719
DB 661 TUNGITSEIGALVAVVGVGCGKSLLSALLAEMDKVGHVAKGSVAVVPOQAMION 720
QY 720 DSIARENIFGQLEEPYRYSVIOCALLPDLTELPSGDRTEIGKGVNLGGQKQVSLA 779
DB 721 DSIARENIFGQLEEPYRYSVIOCALLPDLTELPSGDRTEIGKGVNLGGQKQVSLA 780
QY 780 RAVSNADIVLFDPLSLAVDAVHAKHIFENVIGPKGMKXKTRILVTHMSYLPQVDVVI 839
DB 781 RAVSNADIVLFDPLSLAVDAVHAKHIFENVIGPKGMKXKTRILVTHMSYLPQVDVVI 840
QY 840 VMSGKISEMGSYQELLRDGAFABFVRYTANTEQDASEDSKNGVSGKESKPVENG 900
DB 841 VMSGKISEMGSYQELLRDGAFABFVRYTANTEQDASEDSKNGVSGKESKPVENG 900
QY 900 MLVTPDSAGKQOLRDLSSSSVSQDISRHNSSTALQKAEKKEETWLMREADKQOTGVK 959
DB 901 MLVTPDSAGKQOLRDLSSSSVSQDISRHNSSTALQKAEKKEETWLMREADKQOTGVK 959
QY 959 ILVTVDAVGKPLQRLNLSNSSHVVNTQOHSSTALQKS -GKKEETWLMREADKQOTGVK 959
DB 959 ILVTVDAVGKPLQRLNLSNSSHVVNTQOHSSTALQKS -GKKEETWLMREADKQOTGVK 959
QY 960 LSVYVMDVWKAIGLFIISFIFLPMCNHVSALASNYWMLMTDD -PIVNGTOEHTKVLV 1018
DB 960 LSVYVMDVWKAIGLFIISFIFLPMCNHVSALASNYWMLMTDD -PIVNGTOEHTKVLV 1019
QY 1019 YGALGISQGIAPVGYMAVSTIGLILASRCLHVDLHSLNSPMSFFERTPSGNLVNRP 1078
DB 1020 YGALGISQGIAPVGYMAVSTIGLILASRCLHVDLHSLNSPMSFFERTPSGNLVNRP 1079
QY 1079 ELDTVDMSIPVYIKMFGSLFNVIAGACIVILLAPPIAIIIPRGLIYFFQRYVASSR 1138
DB 1080 ELDTVDMSIPVYIKMFGSLFNVIAGACIVILLAPPIAIIIPRGLIYFFQRYVASSR 1139
QY 1139 QLKRLSEVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLYKVENOKAVYPSIVANRW 1198
DB 1140 QLKRLSEVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLYKVENOKAVYPSIVANRW 1199
QY 1199 LAVRLECVGNICVLFAALFAVISRHSLSAGLVGLSVSISQVTTYLWNLVMSSEMETNI 1258
DB 1200 LAVRLECVGNICVLFAALFAVISRHSLSAGLVGLSVSISQVTTYLWNLVMSSEMETNI 1259
QY 1259 VAVRLEKYSTKEGAEPWQIOETAPPSGWPQVGYVERRNYCLARREDLDFLRHINVTYN 1318
DB 1260 VAVRLEKYSTKEGAEPWQIOETAPPSGWPQVGYVERRNYCLARREDLDFLRHINVTYN 1319
QY 1319 GGEKVGIVGRGAKSSLTGLPFINSAAGEEIIIDGINIKIGLHDLRFKITTIIPODPV 1378
DB 1320 GGEKVGIVGRGAKSSLTGLPFINSAAGEEIIIDGINIKIGLHDLRFKITTIIPODPV 1379
QY 1379 LFGSGLRNKLDLPFSQYSDDEEYVTSLELAHLKD FVSALPDKLDHBCAEGENLVGQRLV 1438
DB 1380 LFGSGLRNKLDLPFSQYSDDEEYVTSLELAHLKD FVSALPDKLDHBCAEGENLVGQRLV 1439
QY 1439 CLARALLRKTILVDEATAVVDLETDLIQSTIRTOGEDCTVLTIAHRLMTINDYTVI 1498
DB 1440 CLARALLRKTILVDEATAVVDLETDLIQSTIRTOGEDCTVLTIAHRLMTINDYTVI 1499

QY 1319 GGEVGIYGTGAGKSSITLGLPRINESAGEEIIIDGINIAKIGHADIRKITTIIIPDDV 1378
DB 1320 GGEVGIYGTGAGKSSITLGLPRINESAGEEIIIDGINIAKIGHADIRKITTIIIPDDV 1379
QY 1379 LFSGLSLWNLDLPFOYSDSEEVWTSLELAHKDFVSALPDKLDHEGAGEGMLVGQRQV 1438
DB 1380 LFPSSLRNLDLPFOYSDSEEVWMALELAHKDFVSALPDKLDHEGAGEGMLVGQRQV 1439
QY 1439 CLARALLRKTKILVLDEATAVLETDLDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVI 1498
DB 1440 CLARALLRKTKILVLDEATAVLETDLDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVI 1499
QY 1499 VLDKGEIQEYCAPSDLLQQRGLFYSMAVDAGLV 1531
DB 1500 VLDKGEIRECAPSELLOQRGVFYSMAVDAGLV 1532
RESULT 14
Q810G9_RAT PRELIMINARY; PRT; 1523 AA.
ID Q810G9_RAT PRELIMINARY; PRT; 1523 AA.
AC Q810G9;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE ATP-binding cassette protein CI variant A.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Takayanagi S., Ishikawa T.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBSJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY174892; AAC44983.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF_2; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1523 AA; 170505 MW; E40337051A1CB9C6 CRC64;
Query Match 69.6%; Score 6892; DB 2; Length 1523;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;

QY 1 MALSGFCAGDSDDPLMDNNTNSNPDTKCFONTVYVWVPCFYLMACFPYFLYLSRH 60
DB 1 MALSGFCAGDSDDPLMDNNTNSNPDTKCFONTVYVWVPCFYLMACFPYFLYLSRH 60
QY 61 DRGYIQMTPLNKTKTALGFLMWICWADLFYSFWERSRGIPLAPVFLVSTLLGITTLIA 120
DB 61 DRGYIQMTPLNKTKTALGFLMWICWADLFYSFWERSRGIPLAPVFLVSTLLGITTLIA 120
QY 121 TFLIOLERRKGVSSGIMLTFWLVALVCALALNRSKIMTALKEDAQVDLFPDITFYFYFS 180
DB 121 TFLIOLERRKGVSSGIMLTFWLVALVCALALNRSKIMTALKEDAQVDLFPDITFYFYFS 180
QY 121 TFLIOLERRKGVSSGIMLTFWLVALVCALALNRSKIMTALKEDAQVDLFPDITFYFYFS 180
DB 121 TFLIOLERRKGVSSGIMLTFWLVALVCALALNRSKIMTALKEDAQVDLFPDITFYFYFS 180
QY 181 LLLIQLVLCGSDRSLPFSSETIHDNCPSSASFLSRITFFWITGLI VRGQROPLEGSD 240
DB 181 LLLIQLVLCGSDRSLPFSSETIHDNCPSSASFLSRITFFWITGLI VRGQROPLEGSD 240
QY 181 LLLIQLVLCGSDRSLPFSSETIHDNCPSSASFLSRITFFWITGLI VRGQROPLEGSD 240
DB 181 LLLIQLVLCGSDRSLPFSSETIHDNCPSSASFLSRITFFWITGLI VRGQROPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVYVQWKKKCACTRKQPVVYVS -SKDPQPKSSSVNDNEVEA 299
DB 241 LMSLNKEDTSEQVVPVYVQWKKKCACTRKQPVVYVS -SKDPQPKSSSVNDNEVEA 299
QY 241 LMSLNKEDTSEQVVPVYVQWKKKCACTRKQPVVYVS -SKDPQPKSSSVNDNEVEA 299
DB 241 LMSLNKEDTSEQVVPVYVQWKKKCACTRKQPVVYVS -SKDPQPKSSSVNDNEVEA 299
QY 300 LMSLNKEDTSEQVVPVYVQWKKKCACTRKQPVVYVS -SKDPQPKSSSVNDNEVEA 300
DB 300 LMSLNKEDTSEQVVPVYVQWKKKCACTRKQPVVYVS -SKDPQPKSSSVNDNEVEA 300
QY 301 LMSLNKEDTSEQVVPVYVQWKKKCACTRKQPVVYVS -SKDPQPKSSSVNDNEVEA 301
DB 301 LMSLNKEDTSEQVVPVYVQWKKKCACTRKQPVVYVS -SKDPQPKSSSVNDNEVEA 301
QY 360 DMQGYFYTVLLFVTACIQLTVAHQYFHCIVSGNRKIKTAVIGAVYRKALVITNSARKST 419
DB 360 DMQGYFYTVLLFVTACIQLTVAHQYFHCIVSGNRKIKTAVIGAVYRKALVITNSARKST 419
QY 361 DMQGYLYTALLFVSAQIQLTVAHQYFHCIVSGNRKIKTAVIGAVYRKALVITNSARKST 420
DB 361 DMQGYLYTALLFVSAQIQLTVAHQYFHCIVSGNRKIKTAVIGAVYRKALVITNSARKST 420
QY 420 VGEIYVNLMSVDAQRFMDLATYINMINSAPLOVITLALYLMWNLGSPVLGAVAVWLVVPV 479
DB 420 VGEIYVNLMSVDAQRFMDLATYINMINSAPLOVITLALYLMWNLGSPVLGAVAVWLVVPV 479
QY 421 VGEIYVNLMSVDAQRFMDLATYINMINSAPLOVITLALYLMWNLGSPVLGAVAVWLVVPV 480
DB 421 VGEIYVNLMSVDAQRFMDLATYINMINSAPLOVITLALYLMWNLGSPVLGAVAVWLVVPV 480
QY 480 NAWAMKTKTYQVAMHMSKDNRIKLMEILNGIKVLKLYAMELAFKXVLAIRQEBELKVL 539
DB 480 NAWAMKTKTYQVAMHMSKDNRIKLMEILNGIKVLKLYAMELAFKXVLAIRQEBELKVL 539
QY 481 NAWAMKTKTYQVAMHMSKDNRIKLMEILNGIKVLKLYAMELAFKXVLAIRQEBELKVL 540
DB 481 NAWAMKTKTYQVAMHMSKDNRIKLMEILNGIKVLKLYAMELAFKXVLAIRQEBELKVL 540
QY 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAFAVSLALEFNILPPLNITL 599
DB 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAFAVSLALEFNILPPLNITL 599
QY 541 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAFAVSLALEFNILPPLNITL 600
DB 541 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAFAVSLALEFNILPPLNITL 600
QY 600 PMVTSIYQASVSLKRLRIFLSHEELPDSIERKPVVDGGGTNSITTRNAFTTARSDPP 659
DB 600 PMVTSIYQASVSLKRLRIFLSHEELPDSIERKPVVDGGGTNSITTRNAFTTARSDPP 659
QY 601 PMVTSIYQASVSLKRLRIFLSHEELPDSIERKPVVDGGGTNSITTRNAFTTARSDPP 660
DB 601 PMVTSIYQASVSLKRLRIFLSHEELPDSIERKPVVDGGGTNSITTRNAFTTARSDPP 660
QY 660 TLNGITFISIPGALVAVVGVGVCGKSSLSLALMEKRVGEVAVIKGSVAVVPOQAMION 719
DB 660 TLNGITFISIPGALVAVVGVGVCGKSSLSLALMEKRVGEVAVIKGSVAVVPOQAMION 719
QY 661 TLNGITFISIPGALVAVVGVGVCGKSSLSLALMEKRVGEVAVIKGSVAVVPOQAMION 720
DB 661 TLNGITFISIPGALVAVVGVGVCGKSSLSLALMEKRVGEVAVIKGSVAVVPOQAMION 720
QY 720 DLSRENILFGCOLPEPYYRYVQAQCALLPDLIELPSGDRTEIGKGVNLSCGQKORVSLA 779
DB 720 DLSRENILFGCOLPEPYYRYVQAQCALLPDLIELPSGDRTEIGKGVNLSCGQKORVSLA 779
QY 721 DLSRENILFGCOLPEPYYRYVQAQCALLPDLIELPSGDRTEIGKGVNLSCGQKORVSLA 780
DB 721 DLSRENILFGCOLPEPYYRYVQAQCALLPDLIELPSGDRTEIGKGVNLSCGQKORVSLA 780
QY 780 RAVYSNADITYLPDPLSAVDAAVGHKIFENVYIGKMKLKKTRILLVTHSMGYLPQVDYII 839
DB 780 RAVYSNADITYLPDPLSAVDAAVGHKIFENVYIGKMKLKKTRILLVTHSMGYLPQVDYII 839
QY 781 RAVYSNADITYLPDPLSAVDAAVGHKIFENVYIGKMKLKKTRILLVTHSMGYLPQVDYII 840
DB 781 RAVYSNADITYLPDPLSAVDAAVGHKIFENVYIGKMKLKKTRILLVTHSMGYLPQVDYII 840
QY 840 VMSGKTSIEMGYSOELLARODAPAEPLRTYASTEOBDAEENGVYGVSGPKKAKOMENG 899
DB 840 VMSGKTSIEMGYSOELLARODAPAEPLRTYASTEOBDAEENGVYGVSGPKKAKOMENG 899
QY 841 VMSGKTSIEMGYSOELLARODAPAEPLRTYASTEOBDAEENGVYGVSGPKKAKOMENG 900
DB 841 VMSGKTSIEMGYSOELLARODAPAEPLRTYASTEOBDAEENGVYGVSGPKKAKOMENG 900
QY 900 MLVYDSAGKQOROLSSSSSYSGDISRHNSHTAEILOKAAKKEBTWKLMEADKAQTGVK 959
DB 900 MLVYDSAGKQOROLSSSSSYSGDISRHNSHTAEILOKAAKKEBTWKLMEADKAQTGVK 959
QY 901 ILVYDAGKPL-----HSVYVNOGHSSTALQGS -GVKEFTWKLMEADKAQTGVK 950
DB 901 ILVYDAGKPL-----HSVYVNOGHSSTALQGS -GVKEFTWKLMEADKAQTGVK 950
QY 960 LSVYWDYKAIQLFISLSTIFLPMCHVSAIASYVWLSLWTD -PIYNGTOEHTKRLSV 1018
DB 960 LSVYWDYKAIQLFISLSTIFLPMCHVSAIASYVWLSLWTD -PIYNGTOEHTKRLSV 1018
QY 951 LSVYWDYKAIQLFISLSTIFLPMCHVSAIASYVWLSLWTD -PIYNGTOEHTKRLSV 1010
DB 951 LSVYWDYKAIQLFISLSTIFLPMCHVSAIASYVWLSLWTD -PIYNGTOEHTKRLSV 1010
QY 1019 YGALGISGIVAPGVSAVSTIGLILASRCVHVDLHSTIRSPMSFEETPSGNTVNRPSK 1078
DB 1019 YGALGISGIVAPGVSAVSTIGLILASRCVHVDLHSTIRSPMSFEETPSGNTVNRPSK 1078
QY 1011 YGALGISGIVAPGVSAVSTIGLILASRCVHVDLHSTIRSPMSFEETPSGNTVNRPSK 1070
DB 1011 YGALGISGIVAPGVSAVSTIGLILASRCVHVDLHSTIRSPMSFEETPSGNTVNRPSK 1070
QY 1079 ELDTVDSMTPEVIAKFMGSLFNVIGACIVILATPIAIIIPPLGLIYFFQRFVASSR 1138
DB 1079 ELDTVDSMTPEVIAKFMGSLFNVIGACIVILATPIAIIIPPLGLIYFFQRFVASSR 1138
QY 1071 ELDTVDSMTPEVIAKFMGSLFNVIGACIVILATPIAIIIPPLGLIYFFQRFVASSR 1130
DB 1071 ELDTVDSMTPEVIAKFMGSLFNVIGACIVILATPIAIIIPPLGLIYFFQRFVASSR 1130
QY 1139 QLRLESVSRSPVYSHNETLLGVSVIRAFEOERFHQSGLKVDENOKAYYPSI VNRW 1198
DB 1139 QLRLESVSRSPVYSHNETLLGVSVIRAFEOERFHQSGLKVDENOKAYYPSI VNRW 1198

Db 1131 QLRLESVSRSPVSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAVYPSIVANRW 1190
QY 1199 LAVRELCVNCIVTFAALFAVISRHSLSAGLVGSVSLQVTTATMLVMSSEMENI 1258
Db 1191 LAVRELCVNCIVTFAALFAVISRHSLSAGLVGSVSLQVTTATMLVMSSEMENI 1250
QY 1259 VAVERLKEYSBTEKEAPWQIOETAPSSMPQGVAFENYCLARYEDLDFLRHINTVIN 1318
Db 1251 VAVERLKEYSBTEKEASWQIOETAPSPCTWPSGRFEFDYCLRRYEDLDVLRKHINTVIE 1310
QY 1319 GGEKVGIVRTGAGKSLTTLGLFRINESAGEEIIIDGINAKIGHMLRFXITII PQDPV 1378
Db 1311 GGEKVGIVRTGAGKSLTTLGLFRINESAGEEIIIDGINAKIGHMLRFXITII PQDPV 1370
QY 1379 LFGSLRWNLDPFGQYSDEEWTSLTSLAHKDFVSALBDKIDHECAEGENLTVGQROLY 1438
Db 1371 LFGSLRWNLDPFGQYSDEEWTSLTSLAHKDFVSALBDKIDHECAEGENLTVGQROLY 1430
QY 1439 CLARALLRKTILVDEATAVDLETDLDLIQSTRQPEDCTVLTIAHRLNTIMDYTRVI 1498
Db 1431 CLARALLRKTILVDEATAVDLETDLDLIQSTRQPEDCTVLTIAHRLNTIMDYTRVI 1490
QY 1499 VLDKGEIQEYGA PSDLLQORGLFYSMAXDAGLV 1531
Db 1491 VLDKGEIREFCAPSELQORGVFYSMAXDAGLV 1523

RESULT 15
Q5F364_CHICK
ID Q5F364_CHICK PRELIMINARY; PRT; 1525 AA.
AC Q5F364;
DT 10-MAY-2005 (Tremblrel. 30. Created)
DT 10-MAY-2005 (Tremblrel. 30. Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30. Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04.32d20;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Plachy J., Kauter S., Blagodatzi A., Kostovska D., Kotler M.,
RA "Full-length cDNAs from chicken bursa lymphocytes to facilitate
RT gene function analysis";
RL Genome Biol. 6:R6-R6(2005).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AJ851786; CAH5420.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:nucleotide binding; IEA.
DR GO; GO:0000166; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR00719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

KW ATP-binding; Hypothetical protein; Nucleotide-binding; Repeat;
KW Transport. 1525 AA; 170972 MW; 3872456711B878FD CR664;
SQ SEQUENCE 1525 AA; 170972 MW; 3872456711B878FD CR664;
Query Match 63.6%; Score 6301; DB 2; Length 1525;
Best Local Similarity 77.6%; Pred. No. 0;
Matches 1193; Conservative 176; Mismatches 148; Indels 20; Gaps 7;
QY 1 MALRGFSGADSDPDLMMWNTMTNSNDFPKCFQNTLVWVPCYTLVACPFYFLYSRH 60
Db 1 MGILSLSDASASEPFWMTNLWHTENPDPFOCFQNTLVWVPCYTLVACPFYFLYSRH 60
QY 61 DRGYIOMTPINKTKTALGFLIMVCMADLFPSFERSRGIFLAVPLVSPFLITTLA 120
Db 61 DRGYIOMSLINKATLGLIMIVCMADLFPSFERSQNTFRAPFLISPLVLTLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFWVALVCAALRSKIMTALKEDAQVDFPRDITFYVFS 180
Db 121 TFLIOLERRKGVSSGIMLTFWVALVCAALRSKIMTALKEDAQVDFPRDITFYVFS 180
QY 181 LLLIQLVLSGSDRSPLFSETIHDNPPCPSSASFLSRITFWITGLIVRGYRQPLGSD 240
Db 181 LLLIQLVLSGSDRSPLFSETIHDNPPCPSSASFLSRITFWITGLIVRGYRQPLGSD 240
QY 241 LLLVQLILSCFPEKPLFSEAVNDPKCPERSASFLSRITFWITGLIMIQHRRLPKAD 240
Db 241 LLLVQLILSCFPEKPLFSEAVNDPKCPERSASFLSRITFWITGLIMIQHRRLPKAD 240
QY 241 LMSLNKEDTSEOVVPLVQWKKCECAKTRKOPVKNVSSKDPQPKXSSKVDAN----EE 296
Db 241 LMSLNKEDTSEIVPGIAKQWAKEMAKTKKOPLMNLTSK----KQKSDSNGEWE 295
QY 297 VEALIVKSPQENNPSEFKVLYKTFPGYFLMSPEFFKAIDHLMESGPOLIKLIFVNDT 356
Db 297 VEALIVKSPQENNPSEFKVLYKTFPGYFLMSPEFFKAIDHLMESGPOLIKLIFVNDT 356
QY 357 KAPWQGFYFTVLFTFACIQTLYLQYFHICVSGKRITKAVYGAAYRAVLVITNSARK 416
Db 357 KAPWQGFYFTVLFTFACIQTLYLQYFHICVSGKRITKAVYGAAYRAVLVITNSARK 416
QY 417 SSTVGEIVNMSVDAQFMDLATYINMISAPLOVTLALVLMNTGPSVLAGAVVLM 476
Db 417 SSTVGEIVNMSVDAQFMDLATYINMISAPLOVTLALVLMNTGPSVLAGAVVLM 476
QY 477 VPVNAVAMAKTKTYQVAMHKSNDNRILAMEILNGIKVLYKLYAMELAFKQVLAIROEL 536
Db 477 VPVNAVAMAKTKTYQVAMHKSNDNRILAMEILNGIKVLYKLYAMELAFKQVLAIROEL 536
QY 537 KVLKKSAYLAVSGFTVTCPPFLVALCTFAVYVITDENNITDAQFASLLENILRPPL 596
Db 537 KVLKKSAYLAVSGFTVTCPPFLVALCTFAVYVITDENNITDAQFASLLENILRPPL 596
QY 597 NILPMVTSIVQASVSLKRLRIFLSHELAPDSIERPVVDGGTNSITVRNMTFQRS 656
Db 597 NILPMVTSIVQASVSLKRLRIFLSHELAPDSIERPVVDGGTNSITVRNMTFQRS 656
QY 657 DPPTLNGITISIEGALVAVAVGVCCKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAW 716
Db 657 DPPTLNGITISIEGALVAVAVGVCCKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAW 716
QY 716 DPPTLNGITISIEGALVAVAVGVCCKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAW 713
Db 716 DPPTLNGITISIEGALVAVAVGVCCKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAW 713
QY 717 IONDSLENTLFGCQLEBPYRSVYQACALLPDLLEIIPSGDRITGEGYNLSGGQKQRY 776
Db 717 IONDSLENTLFGCQLEBPYRSVYQACALLPDLLEIIPSGDRITGEGYNLSGGQKQRY 773
QY 777 SLARAVYSNADYLPDPLSAVDAHCKHIFENVIGKGMKLNKTRILTYHMSYLPQVD 836
Db 777 SLARAVYSNADYLPDPLSAVDAHCKHIFENVIGKGMKLNKTRILTYHMSYLPQVD 833
QY 837 VTIIVSGKISSENGVQELARDAFAEFRTYASTEOQDAEENGVTGVSQP-GKEAKQ 895
Db 837 VTIIVSGKISSENGVQELARDAFAEFRTYASTEOQDAEENGVTGVSQP-GKEAKQ 895
QY 895 VTIIVSGKISSENGVQELARDAFAEFRTYASTEOQDAEENGVTGVSQP-GKEAKQ 895
Db 895 VTIIVSGKISSENGVQELARDAFAEFRTYASTEOQDAEENGVTGVSQP-GKEAKQ 895
QY 954 QTQGVKLVSVMWVKAIGLIFISFLIFLFCNHYVALASVYMLSLMTDPIVNGTQHTK 1013
Db 954 QTQGVKLVSVMWVKAIGLIFISFLIFLFCNHYVALASVYMLSLMTDPIVNGTQHTK 1013

```

Db      948  KTGKAVKATVWEYKKAIGLYISFLSVFLMCNHLASLANSYMLSLMTDDPVVNGTQOYTN 1007
QY      1014  VRLSVYGAIGISOGIAVGYMAVSIIGIILASRCJLHVDLHSLRSFPMSPFEPTPSGNLV 1073
Db      1008  VRLGVYGAIGISOGIAVGYMAVSIIGIIPASRHHLDLHNVLSRSPMSPFEPTPSGNLV 1067
QY      1074  NRSFKELDTVDMSIPEVIMKFMGSLFNVI GACIVILATPIAAIIPPLGLIYFPVQRFY 1133
Db      1068  SRFSKEIDTIDSTIPILKMFMGSTFNVIGACIITLLATPIAAVVIPLGLVYLLVQRFY 1127
QY      1134  VASSRQLKRLSVSRSPVYSHPNETLLGVSVIRAFEEQERFIHQSDLVKVDENOKAYYPSI 1193
Db      1128  VATSRQLKRLSVSRSPVYSHPNETLLGVSVIRAFEEQERFIHQSDLVKVDENOKAYYPSI 1187
QY      1194  VANRMLAVRLBCEVNCIVLPALPAVIRSHSISAGLVGYSYSLOVTTYLNMVLVRMSSE 1253
Db      1188  VANRMLAVRLBCEVNCIVLPALPAVIRSHSISAGLVGYSYSLOVTTYLNMVLVRMTSD 1247
QY      1254  METNIVAVERLKEYSETEKEAPWQIQETAPPSWPQVGRVEFRNYCLARYREDDLPVLRHI 1313
Db      1248  LETNIVAVERLKEYSETEKEAPWQIQETAPPSWPQVGRVEFRNYCLARYREDDLPVLRHI 1307
QY      1314  NVTINGGEKVGIVGRTGAGKSLTGLPRINSAEGEIIIDGINIAKIGLHDLRPKITII 1373
Db      1308  NVTINGGEKVGIVGRTGAGKSLTGLPRINSAEGEIIIDGINIAKIGLHDLRPKITII 1367
QY      1374  PODPVLFGSGLRMNLDPPSOYSDEEWMTSLELAHLKDFVSALPDKLDHECABGGENLSVG 1433
Db      1368  PODPVLFGSGLRMNLDPPSOYSDEEWMTSLELAHLKDFVSALPDKLDHECABGGENLSVG 1427
QY      1434  QROLVCLARALLRKTIIIVLDEATAVLDLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMD 1493
Db      1428  QROLVCLARALLRKTIIIVLDEATAVLDLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMD 1487
QY      1494  YTRVIVLDKGEIQEYGAPSDLIQORGLFYSMAKXAGL 1530
Db      1488  YTRVIVLDKGEIQEYGAPSDLIQORGLFYSMAKXAGL 1524

```

Search completed: December 15, 2005, 15:33:24
 Job time : 257.254 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:25 ; Search time 38.351 Seconds
(without alignments)
4076.547 Million cell updates/sec

Title: US-10-665-283-4

Perfect score: 9734
Sequence: 1 MALRGGFCASGSDPLMDMNV.....NTIKVPRPLCTARQLDSDRS 1691

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/6_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/H_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/R_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	80.7	1531	1 US-08-463-092B-4	Sequence 4, Appli
2	7860	80.7	1531	1 US-08-462-109A-4	Sequence 4, Appli
3	7860	80.7	1531	1 US-08-460-907B-4	Sequence 4, Appli
4	7860	80.7	1531	2 US-08-463-179A-4	Sequence 4, Appli
5	7860	80.7	1531	2 US-08-461-384B-4	Sequence 4, Appli
6	7860	80.7	1531	2 US-09-647-140B-19	Sequence 19, Appli
7	7849	80.6	1531	1 US-08-141-893-2	Sequence 2, Appli
8	7849	80.6	1531	1 US-08-463-092B-2	Sequence 2, Appli
9	7849	80.6	1531	1 US-08-462-109A-2	Sequence 2, Appli
10	7849	80.6	1531	1 US-08-460-907B-2	Sequence 2, Appli
11	7849	80.6	1531	2 US-08-463-179A-2	Sequence 2, Appli
12	7849	80.6	1531	2 US-08-461-384B-2	Sequence 2, Appli
13	7849	80.6	1531	2 US-08-407-207A-2	Sequence 2, Appli
14	7002.5	71.9	1528	1 US-08-463-092B-6	Sequence 6, Appli
15	7002.5	71.9	1528	1 US-08-462-109A-6	Sequence 6, Appli
16	7002.5	71.9	1528	1 US-08-460-907B-6	Sequence 6, Appli
17	7002.5	71.9	1528	2 US-08-463-179A-6	Sequence 6, Appli
18	7002.5	71.9	1528	2 US-08-461-384B-6	Sequence 6, Appli
19	4487.5	46.1	1527	2 US-09-647-140B-6	Sequence 6, Appli
20	4487.5	46.1	1530	2 US-09-647-140B-33	Sequence 33, Appli
21	3391.5	34.8	1503	2 US-09-647-140B-8	Sequence 8, Appli
22	3385.5	34.8	1503	2 US-09-792-616-3	Sequence 3, Appli
23	3218.5	33.1	1498	2 US-09-792-616-9	Sequence 9, Appli
24	2374.5	24.4	1621	2 US-08-972-927-3	Sequence 3, Appli
25	2349.5	24.1	1622	2 US-08-972-927-6	Sequence 6, Appli
26	2266.5	23.3	1325	2 US-09-647-140B-2	Sequence 2, Appli
27	2212.5	22.7	1464	2 US-10-012-896-1008	Sequence 1008, Ap

ALIGNMENTS

28	2195.5	22.6	1261	2 US-09-439-313-538	Sequence 538, App
29	2195.5	22.6	1261	2 US-09-636-215-538	Sequence 538, App
30	2195.5	22.6	1261	2 US-09-685-166A-538	Sequence 538, App
31	2195.5	22.6	1261	2 US-09-679-426-538	Sequence 538, App
32	2195.5	22.6	1261	2 US-09-759-143-538	Sequence 538, App
33	2195.5	22.6	1261	2 US-09-651-236-538	Sequence 538, App
34	2195.5	22.6	1261	2 US-09-657-279-538	Sequence 538, App
35	2195.5	22.6	1261	2 US-10-012-896-538	Sequence 538, App
36	2195.5	22.6	1261	2 US-10-012-896-1009	Sequence 1009, App
37	2118	21.8	1581	2 US-08-726-320-3	Sequence 3, Appli
38	2118	21.8	1581	2 US-09-208-716-3	Sequence 3, Appli
39	2115	21.7	1228	2 US-09-439-313-537	Sequence 537, App
40	2115	21.7	1228	2 US-09-636-215-537	Sequence 537, App
41	2115	21.7	1228	2 US-09-685-166A-537	Sequence 537, App
42	2115	21.7	1228	2 US-09-679-426-537	Sequence 537, App
43	2115	21.7	1228	2 US-09-759-143-537	Sequence 537, App
44	2115	21.7	1228	2 US-09-651-236-537	Sequence 537, App
45	2115	21.7	1228	2 US-09-657-279-537	Sequence 537, App

RESULT 1
US-08-463-092B-4
Sequence 4, Application US/08463092B
Patent No. 5766880
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
NUMBER OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1531 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-463-0928-4

Query Match 80.7%; Score 7860; DB 1; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALRFGSADGSDPLMDMNTWNTSNPDTKCFQNTVLVWYPCFTLWACFPFFLYLSRH 60
DB 1 MALRFGSADGSDPLMDMNTWNTSNPDTKCFQNTVLVWYPCFTLWACFPFFLYLSRH 60
QY 61 DRGIQMTPLNKTATLGEILMTICWADLFYSFMRSRSGIFLAPFLVSPILLGTTLLA 120
DB 61 DRGIQMTPLNKTATLGEILMTICWADLFYSFMRSRSGIFLAPFLVSPILLGTTLLA 120
QY 121 TFLQLERRKGVSSGIMLTFLWVALVICALILRSKIMTALKEDAQVDLFRDITFYVYS 180
DB 121 TFLQLERRKGVSSGIMLTFLWVALVICALILRSKIMTALKEDAQVDLFRDITFYVYS 180
QY 181 LLLIQLVLSGSDSPLESETIHDNPPCESSASFLSRTITWWTGLIVRGYROPLEGSD 240
DB 181 LLLIQLVLSGSDSPLESETIHDNPPCESSASFLSRTITWWTGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPQKMNPSLFEVLVYKTEGPFYFLMSFFPKAIDHLMFSGPQILKLIKFVNDTKAPD 360
DB 301 IVKSPQKMNPSLFEVLVYKTEGPFYFLMSFFPKAIDHLMFSGPQILKLIKFVNDTKAPD 360
QY 361 WQGFYTYVLLVTVACLQTLVTHQYTHICFVSGMRITKAVIGAAYRKALVITNSARKSTV 420
DB 361 WQGFYTYVLLVTVACLQTLVTHQYTHICFVSGMRITKAVIGAAYRKALVITNSARKSTV 420
QY 421 GEIVNMSVDARFMDLATYINMTWSAPLOVTLALYLMINLGPVLAVGVAVMTMAYVN 480
DB 421 GEIVNMSVDARFMDLATYINMTWSAPLOVTLALYLMINLGPVLAVGVAVMTMAYVN 480
QY 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKQKYLAIROBELKYLK 540
DB 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKQKYLAIROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTTDENNILDQAIFVSLALFNILRPLNLTLP 600
DB 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTTDENNILDQAIFVSLALFNILRPLNLTLP 600
QY 601 MVISIVASVSIKRLRIFLSHEBELPDSIERRPVKGDDGNTSIVRNAFTWASDPT 660
DB 601 MVISIVASVSIKRLRIFLSHEBELPDSIERRPVKGDDGNTSIVRNAFTWASDPT 660
QY 661 LAGTIFSIPEGALVAVVGQVCGKSSLSALLAEMDKYEGHVAIKGSVAYVPOQAMIOND 720
DB 661 LAGTIFSIPEGALVAVVGQVCGKSSLSALLAEMDKYEGHVAIKGSVAYVPOQAMIOND 720
QY 721 SLRENIIFGCOLLEBYRVSITQACALLPDLLEILPSGDRTEIGEKVNI.SGGQKQFVSLAR 780
DB 721 SLRENIIFGCOLLEBYRVSITQACALLPDLLEILPSGDRTEIGEKVNI.SGGQKQFVSLAR 780
QY 781 AVYSNADIYLPDDPLSAVDANHGKHI FENVIGPKKMLKTKRILIVTHMSSTYLPQVDVILV 840
DB 781 AVYSNADIYLPDDPLSAVDANHGKHI FENVIGPKKMLKTKRILIVTHMSSTYLPQVDVILV 840
QY 841 MSGKISMGSYOELIARDGAFAEFLRTYASTEOBDAEENGVTGSGPGKEAKOMEGM 900
DB 841 MSGKISMGSYOELIARDGAFAEFLRTYASTEOBDAEENGVTGSGPGKEAKOMEGM 900
QY 901 LVTSAGKOLQROLSSSSSYSGDISRHNSTAELOKAEAKKETWLMADRAQOTGOYVL 960
DB 901 LVTSAGKOLQROLSSSSSYSGDISRHNSTAELOKAEAKKETWLMADRAQOTGOYVL 960

```

```

DB 901 LVTSAGKOLQROLSSSSSYSGDISRHNSTAELOKAEAKKETWLMADRAQOTGOYVL 960
QY 961 SVYWDYMKAIQGLFISFSLIFLMCNHVSALASNYMLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYMKAIQGLFISFSLIFLMCNHVSALASNYMLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISGIAVFGYSMAVSIIGIILASRCLHVDLHLSILRSPMSFPERTPSGNLVNRFSKEL 1080
DB 1021 ALGISGIAVFGYSMAVSIIGIILASRCLHVDLHLSILRSPMSFPERTPSGNLVNRFSKEL 1080
QY 1081 DTVDMSIPEVTKMMSGLFENVIGACIVILATPIAIIIPPLGLIFFVQRFVYASSROL 1140
DB 1081 DTVDMSIPEVTKMMSGLFENVIGACIVILATPIAIIIPPLGLIFFVQRFVYASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDKVDENQKAYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDKVDENQKAYPSIVANRWLA 1200
QY 1201 VRLCEVANCIVLPALFAVISRHSLSAGLVGLSVYSLOVTTYINMLVRRSSMETNIVA 1260
DB 1201 VRLCEVANCIVLPALFAVISRHSLSAGLVGLSVYSLOVTTYINMLVRRSSMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMGIQETAPSSWPQVGRVFRNYCLAREDLDPVLRIHIVTTNGG 1320
DB 1261 VERLKEYSETEKEAPMGIQETAPSSWPQVGRVFRNYCLAREDLDPVLRIHIVTTNGG 1320
QY 1321 EKVGIVRTGAGKSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGLRPMNLDPFSQYSDDEWVTSLELAHKDFVSA.LPKLDHECAEGGENLSVGOROLVCL 1440
DB 1381 SGLRPMNLDPFSQYSDDEWVTSLELAHKDFVSA.LPKLDHECAEGGENLSVGOROLVCL 1440
QY 1441 ARALLRKTILVDEATRAVDLETFDDLIQSTIRPOFEDCTVLTIAHRLNTIMDYTRVVL 1500
DB 1441 ARALLRKTILVDEATRAVDLETFDDLIQSTIRPOFEDCTVLTIAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYMAKXAGLV 1531
DB 1501 DKGEIOEGAPSDLLQORGLFYMAKXAGLV 1531

RESULT 2
US-08-462-109A-4
; Sequence 4, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340

```


FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-4

Query Match 80.7%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDNMVNTMTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYSRH 60
DB 1 MALRGFCSADGSDPLMDNMVNTMTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYSRH 60
QY 61 DRGYIQTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAFVFLVSPLLGITTLLA 120
DB 61 DRGYIQTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAFVFLVSPLLGITTLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMLVYALCALILSKIMTAKEDAOVDLPEDITFYFFS 180
DB 121 TFLIQLERRKGVSSGIMLTFMLVYALCALILSKIMTAKEDAOVDLPEDITFYFFS 180
QY 121 TFLIQLERRKGVSSGIMLTFMLVYALCALILSKIMTAKEDAOVDLPEDITFYFFS 180
DB 121 TFLIQLERRKGVSSGIMLTFMLVYALCALILSKIMTAKEDAOVDLPEDITFYFFS 180
QY 181 LLLIQLVLSGSDSPFLFSETHDNPNCPESSASFLSRITFMWITGLIVRGYRPLGSD 240
DB 181 LLLIQLVLSGSDSPFLFSETHDNPNCPESSASFLSRITFMWITGLIVRGYRPLGSD 240
QY 241 LMSLNKEDTSEQVAVLVKWKKECAKTRKQPVKVVYSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVAVLVKWKKECAKTRKQPVKVVYSSKDPAPKSSKVDANEVEAL 300
QY 301 IVSPPOKEMNPSEKFKVYKTFGPFYFLMSFFPKAIDHLMFSGPOLLKILKIFVNDTRAPD 360
DB 301 IVSPPOKEMNPSEKFKVYKTFGPFYFLMSFFPKAIDHLMFSGPOLLKILKIFVNDTRAPD 360
QY 361 WOGFYFVLLFVTACLOTLVHOFYHICFVSGMRKIKTAVIGAVYRKALVITNSARKSSTV 420
DB 361 WOGFYFVLLFVTACLOTLVHOFYHICFVSGMRKIKTAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAQRPMDLATYINMIMSAPLOVITLALVYLLMLTGPSVLAVVWVLPVN 480
DB 421 GEIYNLMSVDAQRPMDLATYINMIMSAPLOVITLALVYLLMLTGPSVLAVVWVLPVN 480
QY 481 AVNAAMKTKTQVAMKSKDNRIKLMNEILNGIKTLKIKYAMELAFKDYLAIRBELVYLK 540
DB 481 AVNAAMKTKTQVAMKSKDNRIKLMNEILNGIKTLKIKYAMELAFKDYLAIRBELVYLK 540
QY 541 KSAVLASVGTFTWCTPELVALCTEFAVYVITDENNNIIDAQPAFVSLALFNILRPLILP 600
DB 541 KSAVLASVGTFTWCTPELVALCTEFAVYVITDENNNIIDAQPAFVSLALFNILRPLILP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKOGGINSITVNNATFTWARSPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKOGGINSITVNNATFTWARSPT 660
QY 661 LNTGTFSPGALVAVVGVCGGSSLSLALAMDVEGHVAKGSVAVYPOQAMQND 720
DB 661 LNTGTFSPGALVAVVGVCGGSSLSLALAMDVEGHVAKGSVAVYPOQAMQND 720
QY 721 SLRENILFGQLEBPYRSVIOACALLPDEILPSGRTEIGEGVNLSSGQKQVSLAR 780

DB 721 SLRENILFGQLEBPYRSVIOACALLPDEILPSGRTEIGEGVNLSSGQKQVSLAR 780
QY 781 AVYSNADIIYFPDPLSADVDAHVGKHFENYIGPFGMLKAKTRILVYTHSMGLPOVDYIIV 840
DB 781 AVYSNADIIYFPDPLSADVDAHVGKHFENYIGPFGMLKAKTRILVYTHSMGLPOVDYIIV 840
QY 841 MSGGKISEMSYOELARDGAPAEFLRTVASTBOEADAEENGVTGVSQPGKEAOMENG 900
DB 841 MSGGKISEMSYOELARDGAPAEFLRTVASTBOEADAEENGVTGVSQPGKEAOMENG 900
QY 901 LVYDSACKOLQOROLSSSSSSGDISRHNHSTABLOKAEKETWKMMEADKAOTGVYL 960
DB 901 LVYDSACKOLQOROLSSSSSSGDISRHNHSTABLOKAEKETWKMMEADKAOTGVYL 960
QY 961 SVYWDYKAIQGLFISFLSIFLFCMNHVSALASNYWLSMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYKAIQGLFISFLSIFLFCMNHVSALASNYWLSMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGYMAVSIQILASRCLHVDLHLSIRSPMSFEKTPSGNLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYMAVSIQILASRCLHVDLHLSIRSPMSFEKTPSGNLVNRFSKEL 1080
QY 1081 DTVDSMIPVYIKMFMSGLFNVIGACIYILATPIAIIIPPLGIFYFVORFYAASSROL 1140
DB 1081 DTVDSMIPVYIKMFMSGLFNVIGACIYILATPIAIIIPPLGIFYFVORFYAASSROL 1140
QY 1141 KRLSVSRSPVYSHFNELLGVSIVYRAFEQBERFIHQSDLKVDENQAYVPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNELLGVSIVYRAFEQBERFIHQSDLKVDENQAYVPSIVANRWLA 1200
QY 1201 VRLCVCNCTVLPALFAVISRSLSGVLGVSYSLOVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLCVCNCTVLPALFAVISRSLSGVLGVSYSLOVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFEFRNYCYRAREDDLPVLRHINVTNG 1320
DB 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFEFRNYCYRAREDDLPVLRHINVTNG 1320
QY 1321 EKVIGVGTAGKSSLLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVIGVGTAGKSSLLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGLSRMLDPPSOYSDEEFTSLBLAKPVPVGLPKLDHECAEGEENISVGOROLVCL 1440
DB 1381 SGLSRMLDPPSOYSDEEFTSLBLAKPVPVGLPKLDHECAEGEENISVGOROLVCL 1440
QY 1441 ABALLRKTILVDEATAVADLETDLIQSTIRTOFEDCTVLTAHRLNTIMDYTRIYVL 1500
DB 1441 ABALLRKTILVDEATAVADLETDLIQSTIRTOFEDCTVLTAHRLNTIMDYTRIYVL 1500
QY 1501 DKGEIOEGYAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEGYAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 3
US-08-460-907B-4
Sequence 4, Application US/08460907B
Patent No. 5891724
GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
RESISTANCE ON A CELL
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: PARTIQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-4

Query Match 80.7%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADSDPDLMDNNTWNTSNPDFTKCFQNTVLWVPCFYLMACFPFYFLYSRH 60
DB 1 MALRGFCADSDPDLMDNNTWNTSNPDFTKCFQNTVLWVPCFYLMACFPFYFLYSRH 60
QY 61 DRGYIOMTPLNKTKTALGFLIMVCMADLFYSFWRSGIPLAPVFLVSPILLGITTLA 120
DB 61 DRGYIOMTPLNKTKTALGFLIMVCMADLFYSFWRSGIPLAPVFLVSPILLGITTLA 120
QY 121 TFLIOLERRKVOSSGIMLTFWLVALVCAALILRSKINTALKEDQVDFDITFYVFS 180
DB 121 TFLIOLERRKVOSSGIMLTFWLVALVCAALILRSKINTALKEDQVDFDITFYVFS 180
QY 181 LLLIOLVLSGSDSPLESETIHDNPPCESASFLSRITFWMTIGLIVRGYROPLEGGSD 240
DB 181 LLLIOLVLSGSDSPLESETIHDNPPCESASFLSRITFWMTIGLIVRGYROPLEGGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPADQPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPADQPKSSKYDANEVEAL 300
QY 301 IVKSPQKEMWNSLFEVLYKTGPFLYMSFPFKAIHDLMMFSGPQILKLILFVNDTKAPD 360
DB 301 IVKSPQKEMWNSLFEVLYKTGPFLYMSFPFKAIHDLMMFSGPQILKLILFVNDTKAPD 360
QY 361 WQGFYTYTLFETVACLOTVLVHOVPHICFVSGMRITKTAIVIGAVYRKALVITNSAKKSTV 420
DB 361 WQGFYTYTLFETVACLOTVLVHOVPHICFVSGMRITKTAIVIGAVYRKALVITNSAKKSTV 420
QY 421 GEIVNLSVDAQRFDLATYINIMWSAPLOVILALYLLMLNLGSPVLGAVAVMVLMEVFN 480

DB 421 GEIVNLSVDAQRFDLATYINIMWSAPLOVILALYLLMLNLGSPVLGAVAVMVLMEVFN 480
QY 481 AVNMMKTKTYQVAHMKSKNRKIKLMEIILNGIKVLKLYAMELAFKDKVLAIRQBELKYLK 540
DB 481 AVNMMKTKTYQVAHMKSKNRKIKLMEIILNGIKVLKLYAMELAFKDKVLAIRQBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVYITDENNIIDAQTAFAVSALFNILRPNILP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVYITDENNIIDAQTAFAVSALFNILRPNILP 600
QY 601 MVISIIYQASVSLKRLRILFISHELEPDSIERRVKGGGNTSTVNAATFTARSDPPT 660
DB 601 MVISIIYQASVSLKRLRILFISHELEPDSIERRVKGGGNTSTVNAATFTARSDPPT 660
QY 661 LMGITPSIPGALVAVVGVQCGKSSILSLALMEADKYEGLVAIKGSVAAYVPOQAWIOND 720
DB 661 LMGITPSIPGALVAVVGVQCGKSSILSLALMEADKYEGLVAIKGSVAAYVPOQAWIOND 720
QY 721 SLRENILFGQLEBPYRSYIQAACALLPDLLEILPSGDRTEIGKGVNLSGQOKORVSLAR 780
DB 721 SLRENILFGQLEBPYRSYIQAACALLPDLLEILPSGDRTEIGKGVNLSGQOKORVSLAR 780
QY 781 AVYSNADIYLPDDPLSAVDHVGKHIPEENVIGPKMKLKNKRIILVTHSMSTLPOVDYIIV 840
DB 781 AVYSNADIYLPDDPLSAVDHVGKHIPEENVIGPKMKLKNKRIILVTHSMSTLPOVDYIIV 840
QY 841 MSGGKISEMSYQELLRDGAFAEFLRTYASTBOBDAEENGVTGVSFGPKGAOKOMENG 900
DB 841 MSGGKISEMSYQELLRDGAFAEFLRTYASTBOBDAEENGVTGVSFGPKGAOKOMENG 900
QY 901 LVYDSAGKOLORQULSSSSSYSGDISRRHNSTABLOKAERKERTWKLMEADKAOTGOVKL 960
DB 901 LVYDSAGKOLORQULSSSSSYSGDISRRHNSTABLOKAERKERTWKLMEADKAOTGOVKL 960
QY 961 SVYWDYMKAIQGLFISFISIFLMCNHVSALASNWLMLMTDDPIVNGTOBHTKRLSYVG 1020
DB 961 SVYWDYMKAIQGLFISFISIFLMCNHVSALASNWLMLMTDDPIVNGTOBHTKRLSYVG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCIADVLDSILSRPMSFERTPSGNLVNRFSEL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCIADVLDSILSRPMSFERTPSGNLVNRFSEL 1080
QY 1081 DTVDMSIPEVIKMGMSLFENVIGACIVILATPIAIIIPPLGIYFFVQGFYVASSROL 1140
DB 1081 DTVDMSIPEVIKMGMSLFENVIGACIVILATPIAIIIPPLGIYFFVQGFYVASSROL 1140
QY 1141 KRLSVSRSPYVSHFNETLIGVSYIRAFEEQERFIHOSDLKVDENOKAYPYSIVANRWLA 1200
DB 1141 KRLSVSRSPYVSHFNETLIGVSYIRAFEEQERFIHOSDLKVDENOKAYPYSIVANRWLA 1200
QY 1201 VRLBCVGNCTVLPALFAVISRHSLSAGLVGSYSIQVTTYVNLVRRMSSEMETNIVA 1260
DB 1201 VRLBCVGNCTVLPALFAVISRHSLSAGLVGSYSIQVTTYVNLVRRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSSWPQVGRVFRNYCYLRYREDLPVLVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPPSSWPQVGRVFRNYCYLRYREDLPVLVLRHINVTINGG 1320
QY 1321 EKVGIIGRTGAKSLLTGLFRINESABGEIILIDGINIAKLGHLDRFKITIIIPQDPVLF 1380
DB 1321 EKVGIIGRTGAKSLLTGLFRINESABGEIILIDGINIAKLGHLDRFKITIIIPQDPVLF 1380
QY 1381 SCSLRMNLDPSSQVSEDEWVTSLELAHKDVSALPDLDBECABGGNLSVGQOLVCL 1440
DB 1381 SCSLRMNLDPSSQVSEDEWVTSLELAHKDVSALPDLDBECABGGNLSVGQOLVCL 1440
QY 1441 ARALLRKTIIIVLEBATAVDELTDLLIGSTIRFOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ARALLRKTIIIVLEBATAVDELTDLLIGSTIRFOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531

Db 1501 DKGEIOEXGAPSDLLQQRGLFYSMAGLAV 1531

RESULT 4
US-08-463-179A-4
Sequence 4, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Coley, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deconti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-4

Query Match 80.7%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFSSAGSGDPLMNMNTTNTSNDPTKCFONTLVVWPCYYLMAACEPFYLYSRH 60
Db 1 MALRGFSSAGSGDPLMNMNTTNTSNDPTKCFONTLVVWPCYYLMAACEPFYLYSRH 60

QY 61 DRGVIQNTPLNKTALGFLMTVCMADELFSFWERSRGIFLAVFVSPFLGITTLLA 120
Db 61 DRGVIQNTPLNKTALGFLMTVCMADELFSFWERSRGIFLAVFVSPFLGITTLLA 120

QY 121 TFLIQLERRRGVSSGIMLTFWVALVCAALILRSKIMTALKEBDAQVDLFRDITFYVYFS 180
Db 121 TFLIQLERRRGVSSGIMLTFWVALVCAALILRSKIMTALKEBDAQVDLFRDITFYVYFS 180

QY 181 LLILQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWMTGLIVNGYRPLEGSD 240
Db 181 LLILQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWMTGLIVNGYRPLEGSD 240

QY 241 LMSLNKEDTSEQVAVLVKMWKKECAKTRKQPVKVVYSSKDPAQPKSSKVDANEVEAL 300

Db 241 LMSLNKEDTSEQVAVLVKMWKKECAKTRKQPVKVVYSSKDPAQPKSSKVDANEVEAL 300

QY 301 IVKSPQKSNPSPLEKVLKTFPGPYFLMSFFPKAIDHLMSSGPOILKLLIKFVNDTAPD 360
Db 301 IVKSPQKSNPSPLEKVLKTFPGPYFLMSFFPKAIDHLMSSGPOILKLLIKFVNDTAPD 360

QY 361 MGGFYTVLLFVTAQCTLVLAHOYFHI CFSQGM IKTAVIGAVYRKALVTNSARKSTV 420
Db 361 MGGFYTVLLFVTAQCTLVLAHOYFHI CFSQGM IKTAVIGAVYRKALVTNSARKSTV 420

QY 421 GEIVNLMSVDAQREMDLATYIMMWSAPLOVITLALYLMNLGSPVLAVGAVMTLAMPVN 480
Db 421 GEIVNLMSVDAQREMDLATYIMMWSAPLOVITLALYLMNLGSPVLAVGAVMTLAMPVN 480

QY 481 AVNANKTKTYQVAMKSKNRIRIKLMEILNGIKVLYKAYAMELAFKDYALAIROBELKYLK 540
Db 481 AVNANKTKTYQVAMKSKNRIRIKLMEILNGIKVLYKAYAMELAFKDYALAIROBELKYLK 540

QY 541 KSAVLASVGTPTWCTPELVALCTFAVYVITDENNILLDAQTAQVSLAFNITLRPNTLP 600
Db 541 KSAVLASVGTPTWCTPELVALCTFAVYVITDENNILLDAQTAQVSLAFNITLRPNTLP 600

QY 601 MVISSIVQASVSLKRLIFLSHEELBPSIERRPVKDGGTNSITVNAFTMARSDPT 660
Db 601 MVISSIVQASVSLKRLIFLSHEELBPSIERRPVKDGGTNSITVNAFTMARSDPT 660

QY 661 LMGITFSPREGALVAVVGQVCGKSLSLALLAEMDRVGHVAIKGSVAVYVPOQAWI QND 720
Db 661 LMGITFSPREGALVAVVGQVCGKSLSLALLAEMDRVGHVAIKGSVAVYVPOQAWI QND 720

QY 721 SLRENILFGQLEBPYRSVIOACALLPDEILPESGRTETGEGVNLSSGQKRVSLAR 780
Db 721 SLRENILFGQLEBPYRSVIOACALLPDEILPESGRTETGEGVNLSSGQKRVSLAR 780

QY 781 AVYSNADITYLPDDPLASVDAVGHKIFENYIGPKMKKNTRIIVTHSMSTLPQVDYIIV 840
Db 781 AVYSNADITYLPDDPLASVDAVGHKIFENYIGPKMKKNTRIIVTHSMSTLPQVDYIIV 840

QY 841 MSGGKISEMWSYQELLARDGAFAEFLRTVASTBOEQAEBENGVTGVSQPKGKAKOMENG 900
Db 841 MSGGKISEMWSYQELLARDGAFAEFLRTVASTBOEQAEBENGVTGVSQPKGKAKOMENG 900

QY 901 LYTDSACKOIORQLSSSSSYSGDISRRHNSTAELQKAKAKEETWKLMEDAKQOTGOVKL 960
Db 901 LYTDSACKOIORQLSSSSSYSGDISRRHNSTAELQKAKAKEETWKLMEDAKQOTGOVKL 960

QY 961 SVYWDYKAIQGLFISFLIFLPMCNHVSALASNYMWSLMTDDPIVNGTOEHTKRLSVYG 1020
Db 961 SVYWDYKAIQGLFISFLIFLPMCNHVSALASNYMWSLMTDDPIVNGTOEHTKRLSVYG 1020

QY 1021 ALGISOGIAVFGYSMAVISGILLASRCLAHVDLHSLIRSPMSFBERPQNLVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYSMAVISGILLASRCLAHVDLHSLIRSPMSFBERPQNLVNRFSKEL 1080

QY 1081 DTVDSTMPEVIKMPMSLENVIGACIVTLATPAIILIPPLGIYFFVORFYVASSROL 1140
Db 1081 DTVDSTMPEVIKMPMSLENVIGACIVTLATPAIILIPPLGIYFFVORFYVASSROL 1140

QY 1141 KRLSVSRSPVYSHPNETLLGVSIVIRAFEOEKPFIHOSDLKVDENQAXYPSIYANWLA 1200
Db 1141 KRLSVSRSPVYSHPNETLLGVSIVIRAFEOEKPFIHOSDLKVDENQAXYPSIYANWLA 1200

QY 1201 VRLCVCNCIVLPALFAVISRSLSGVGLSVSYSLQVTTYLNMVLVRMSSEMETNIVA 1260
Db 1201 VRLCVCNCIVLPALFAVISRSLSGVGLSVSYSLQVTTYLNMVLVRMSSEMETNIVA 1260

QY 1261 VERLKEVSETEKAPMOIQETAPSSWPQGRVFRNYCLARYEDDLDFVLRHINVTNGG 1320
Db 1261 VERLKEVSETEKAPMOIQETAPSSWPQGRVFRNYCLARYEDDLDFVLRHINVTNGG 1320

QY 1321 EKVGIVGRTGAGKSLTILGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380

Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTTITIPDDPVLF 1380
Qy 1381 SGLSRMLNDPFSQYSDSEEWTSLELAHLKDFVSALPKLDHECAEGENLSVGORQVCL 1440
Db 1381 SGLSRMLNDPFSQYSDSEEWTSLELAHLKDFVSALPKLDHECAEGENLSVGORQVCL 1440
Qy 1441 ARALLRKTKIIVLDBATNAVDLETFDDLIOSTIRTOFEDCTVLTAAHRLNTTMDYTRVYL 1500
Db 1441 ARALLRKTKIIVLDBATNAVDLETFDDLIOSTIRTOFEDCTVLTAAHRLNTTMDYTRVYL 1500
Qy 1501 DKGEIOEYGAESPDLLOQRGLFYSMAKDGIV 1531
Db 1501 DKGEIOEYGAESPDLLOQRGLFYSMAKDGIV 1531

RESULT 5

US-08-461-384B-4

Sequence 4, Application US/08461384B

Patent No. 6025473

GENERAL INFORMATION:

APPLICANT: Cole, Susan P.C.

APPLICANT: Dealey, Roger G.

TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESSES:

ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS

STREET: Queen's University at Kingston

CITY: Kingston

STATE: Ontario

COUNTRY: CANADA

ZIP: K7L 3N6

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/461,384B

FILING DATE: 05-JUN-95

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 07/966,923

FILING DATE: 27-OCT-1992

APPLICATION NUMBER: 08/029,340

FILING DATE: 8-MAR-1993

APPLICATION NUMBER: 08/141,893

FILING DATE: 26-OCT-1993

APPLICATION NUMBER: 08/407,207

FILING DATE: 20-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Steeg, Carol Miernicki

REGISTRATION NUMBER: 39,539

REFERENCE/DOCKET NUMBER: 01547

TELECOMMUNICATION INFORMATION:

TELEPHONE: (613) 545-2342

TELEFAX: (613) 545-6853

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1531 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-461-384B-4

Query Match 80.7%; Score 7860; DB 2; Length 1531;

Best local Similarity 100.0%; Pred. No. 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRFGCSADSGDPLMDNMVNTNSPDTKCFONTVLVWVPCFYLMACFPFYLYLSRH 60
Db 1 MALRFGCSADSGDPLMDNMVNTNSPDTKCFONTVLVWVPCFYLMACFPFYLYLSRH 60
Qy 61 DRGIYQMTPLNKTALGFLLMTVCWADLFYSFWERSRGIFLAPVPLVSPTLIGITTLIA 120
Db 61 DRGIYQMTPLNKTALGFLLMTVCWADLFYSFWERSRGIFLAPVPLVSPTLIGITTLIA 120

Db 61 DRGIYQMTPLNKTALGFLLMTVCWADLFYSFWERSRGIFLAPVPLVSPTLIGITTLIA 120
Qy 121 TFLIOLERRKGVSSGIMLTFMWLVVALCALAIIKSKIMTALKEADAQVDLFEDITFYVYFS 180
Db 121 TFLIOLERRKGVSSGIMLTFMWLVVALCALAIIKSKIMTALKEADAQVDLFEDITFYVYFS 180
Qy 181 LLLIQLVLSGCSDFSPLFSETIHDNPNCPRESSASFISRTFWATTGLIVRGYRQPLBESD 240
Db 181 LLLIQLVLSGCSDFSPLFSETIHDNPNCPRESSASFISRTFWATTGLIVRGYRQPLBESD 240
Qy 241 LMSINKEDTSQVAVPVLYKMKKECAKTRKQPVVYVSSKDPAPCKESSKVDANEVEVAL 300
Db 241 LMSINKEDTSQVAVPVLYKMKKECAKTRKQPVVYVSSKDPAPCKESSKVDANEVEVAL 300
Qy 301 IVKSPQKEMNPSLFKVLTKTGPYFLMSFPFKAIDHLMESGPQILKLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFKVLTKTGPYFLMSFPFKAIDHLMESGPQILKLIKFVNDTKAPD 360
Qy 361 WQGYFYTVLFTVTAQLOTLVHGFHICFVSQGMKTAIVGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFYTVLFTVTAQLOTLVHGFHICFVSQGMKTAIVGAVYRKALVITNSARKSSTV 420
Qy 421 GEIVNLSVDAQRPMDLATYINMIMSAPLQVITLALYLLMLNLSVLAGVAVMLVNVN 480
Db 421 GEIVNLSVDAQRPMDLATYINMIMSAPLQVITLALYLLMLNLSVLAGVAVMLVNVN 480
Qy 481 AVMAKTKTYOAVAMKSKDNRIKLMNBIINGIKVLKYAMELAFKDYALROBELKYLK 540
Db 481 AVMAKTKTYOAVAMKSKDNRIKLMNBIINGIKVLKYAMELAFKDYALROBELKYLK 540
Qy 541 KSAVLSAVGTFTWCTPPLVALCTPFAVYVITDENNIIDAOCTAPFSLAFNLRPPNLTLP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTPFAVYVITDENNIIDAOCTAPFSLAFNLRPPNLTLP 600
Qy 601 MVISIVQASVSLKRIPLFSHEELBPSIRRPVYKQGGNSTVRAATFTMARSDPT 660
Db 601 MVISIVQASVSLKRIPLFSHEELBPSIRRPVYKQGGNSTVRAATFTMARSDPT 660
Qy 661 LMGITFSIPBEGALVAVVQVCGKSSLSALLAEMDKVEGHVAIKGSVAVYVPOQAMIOND 720
Db 661 LMGITFSIPBEGALVAVVQVCGKSSLSALLAEMDKVEGHVAIKGSVAVYVPOQAMIOND 720
Qy 721 SLRENIIFGCOLLEBPYRSVIOACALLPDLIELPSGDRTEIGEKGVNLSGQKORVSLAR 780
Db 721 SLRENIIFGCOLLEBPYRSVIOACALLPDLIELPSGDRTEIGEKGVNLSGQKORVSLAR 780
Qy 781 AVYSNADIVLEDDPLSADAHVGHIFENVIGPKMLKNKTRILVTHSMSTLPQVDVITV 840
Db 781 AVYSNADIVLEDDPLSADAHVGHIFENVIGPKMLKNKTRILVTHSMSTLPQVDVITV 840
Qy 841 MSGKISEMGSYOELLARADGAFAEFLRTYASTBOBDAEENGVTGVSQPGKEAKOMENG 900
Db 841 MSGKISEMGSYOELLARADGAFAEFLRTYASTBOBDAEENGVTGVSQPGKEAKOMENG 900
Qy 901 LVYDSAGKOLORQLSSSSSYSGDISRHNSYAEIOLKAEKKEETWKLMEADYAOIGQVYL 960
Db 901 LVYDSAGKOLORQLSSSSSYSGDISRHNSYAEIOLKAEKKEETWKLMEADYAOIGQVYL 960
Qy 961 SVYWDYMKALGLFISFLSIFLMCHNVSAALSNVWLSMTDPIVNGOETHKVALSYVG 1020
Db 961 SVYWDYMKALGLFISFLSIFLMCHNVSAALSNVWLSMTDPIVNGOETHKVALSYVG 1020
Qy 1021 ALGISOGIAVFGYSMAVSIIGILASRCHVDLHLSILSPMSFPERTSGNULVNFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGILASRCHVDLHLSILSPMSFPERTSGNULVNFSEKL 1080
Qy 1081 DTVDSMIDEVIMKFWGSLFNVIAGACIVILATPIAIIIPGLGIYFFVQRPYVASSQOL 1140
Db 1081 DTVDSMIDEVIMKFWGSLFNVIAGACIVILATPIAIIIPGLGIYFFVQRPYVASSQOL 1140
Qy 1141 KRLESVNSPYYSHNEFTLLGVSVYRAFEDEBFTHOSDLVDENOKKAYYSIVANRWLA 1200
Db 1141 KRLESVNSPYYSHNEFTLLGVSVYRAFEDEBFTHOSDLVDENOKKAYYSIVANRWLA 1200

QY 1201 VRLCEVGNCTVLFALFAVSRHSLSAGLVGVSYSLOVTTYLNLVVRSSSEMTNIVA 1260
DB 1201 VRLCEVGNCTVLFALFAVSRHSLSAGLVGVSYSLOVTTYLNLVVRSSSEMTNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNVCYRREDLDFVLRHINTVINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNVCYRREDLDFVLRHINTVINGG 1320
QY 1321 EKVIGVGTGAGKSLTGLFRINESAGEIITIDGINIAKIGLHDLRFKTIITIPQDPVLF 1380
DB 1321 EKVIGVGTGAGKSLTGLFRINESAGEIITIDGINIAKIGLHDLRFKTIITIPQDPVLF 1380
QY 1381 SGSRPMNLDPFSQSDSEEWTSLSLAHLKDFVSALPDGLDHECAGGEMLSVGOQVACL 1440
DB 1381 SGSRPMNLDPFSQSDSEEWTSLSLAHLKDFVSALPDGLDHECAGGEMLSVGOQVACL 1440
QY 1441 ABALLRKTKIIVLDEATAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500
DB 1441 ABALLRKTKIIVLDEATAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500
QY 1501 DKGEIOBYGAPSDLLQORGLFYMAKDAGLV 1531
DB 1501 DKGEIOBYGAPSDLLQORGLFYMAKDAGLV 1531

RESULT 6
US-09-647-140B-19
; Sequence 19, Application US/09647140B
; Patent No. 6803184
; GENERAL INFORMATION:
; APPLICANT: Fox Chase Cancer Center
; APPLICANT: Krush, Gary D.
; APPLICANT: Lee, Kun
; APPLICANT: Belinsky, Martin G.
; APPLICANT: Bain, Lisa J.
; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use Thereof
; FILE REFERENCE: PCCC 98-02
; CURRENT APPLICATION NUMBER: US/09/647, 140B
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US99/06644
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/079, 759
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/095, 153
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-647-140B-19

Query Match 80.7%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVNTSNPDTKCFONTLVVWPCFLYMACFPFFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNTVNTSNPDTKCFONTLVVWPCFLYMACFPFFLYLSRH 60
QY 61 DRGTYQMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAVFVLSPTLLGITTLLA 120
DB 61 DRGTYQMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAVFVLSPTLLGITTLLA 120
QY 121 TPLIQLEERRKGVSSGIMLTFWLVALYCALAILRSKIMTALXBAQVDLPRDITFYVYS 180
DB 121 TPLIQLEERRKGVSSGIMLTFWLVALYCALAILRSKIMTALXBAQVDLPRDITFYVYS 180
QY 181 LLLIQVLSCFSDRSPLFSETHIDPNCPBSSASFLSRITFWWITGLIVGRQPLEGSD 240
DB 181 LLLIQVLSCFSDRSPLFSETHIDPNCPBSSASFLSRITFWWITGLIVGRQPLEGSD 240

DB 181 LLLIQVLSCFSDRSPLFSETHIDPNCPBSSASFLSRITFWWITGLIVGRQPLEGSD 240
QY 241 LMSLNKEDTSGQVVPVYKWKKECATRRKQPVYVYSSKQDPAOPKSSKTDANEVEAL 300
DB 241 LMSLNKEDTSGQVVPVYKWKKECATRRKQPVYVYSSKQDPAOPKSSKTDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKYVYKPGPYFLMSFFPKALHDMMSGPQILKLLIKFVNDTPAD 360
DB 301 IVKSPQKEMNPSLFKYVYKPGPYFLMSFFPKALHDMMSGPQILKLLIKFVNDTPAD 360
QY 361 WQGYFTVLLFVTAQOTLVHLQYFHHICFVSGMRKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGYFTVLLFVTAQOTLVHLQYFHHICFVSGMRKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATTYINMISAPLOYITALYLLMNLGFSVLAGAVMYLVMPVN 480
DB 421 GEIVNLMSVDAQRFMDLATTYINMISAPLOYITALYLLMNLGFSVLAGAVMYLVMPVN 480
QY 481 AVMAKTKTYQVAMKSKDNRIKLMNEILNGIKVLYAMELAKDVLAIROBELKVLK 540
DB 481 AVMAKTKTYQVAMKSKDNRIKLMNEILNGIKVLYAMELAKDVLAIROBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVYTIIDENNIIDAQTAFFVSLALFNILRPPLILP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVYTIIDENNIIDAQTAFFVSLALFNILRPPLILP 600
QY 601 MVISSIVQASVSLKRLIFLSHELEBDSIERRPVKDGGGTSITVNAATFTWARSDEPT 660
DB 601 MVISSIVQASVSLKRLIFLSHELEBDSIERRPVKDGGGTSITVNAATFTWARSDEPT 660
QY 661 LNCITTEIPBGALVAVVGVGCCGKSSLSLALMEMDVBEBHVAIKGSVAVYPOQAWIQND 720
DB 661 LNCITTEIPBGALVAVVGVGCCGKSSLSLALMEMDVBEBHVAIKGSVAVYPOQAWIQND 720
QY 721 SLRENILFGCLSEPPYRSVIOACALLPDLEILPSEGRTEIGEGVNLSSGQKORVSLAR 780
DB 721 SLRENILFGCLSEPPYRSVIOACALLPDLEILPSEGRTEIGEGVNLSSGQKORVSLAR 780
QY 781 AYSNADIVLPDPLSAVDAHVGHKILENVIIGPKMLKNTKRIILVTHSMSVLPQVDYIIV 840
DB 781 AYSNADIVLPDPLSAVDAHVGHKILENVIIGPKMLKNTKRIILVTHSMSVLPQVDYIIV 840
QY 841 MSGGKISEMKSXOELLARDAPAEPLRTAVSTQEOBDAENGVTVSGSPKKAQOMNGM 900
DB 841 MSGGKISEMKSXOELLARDAPAEPLRTAVSTQEOBDAENGVTVSGSPKKAQOMNGM 900
QY 901 LVYDSAGKQORQOLSSSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKXQTOGVKL 960
DB 901 LVYDSAGKQORQOLSSSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKXQTOGVKL 960
QY 961 SVYWDYMKALGLFISFLSIFLFCMCHVVSALASNYMLSLMTDDPIVNGTOBHTKVRSLVYG 1020
DB 961 SVYWDYMKALGLFISFLSIFLFCMCHVVSALASNYMLSLMTDDPIVNGTOBHTKVRSLVYG 1020
QY 1021 ALGISQCIANFGYSMAVSTIGILLASRCLHVDLHSLIRSPMSFFERTPSGNLVNRFSEKL 1080
DB 1021 ALGISQCIANFGYSMAVSTIGILLASRCLHVDLHSLIRSPMSFFERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMIPETVYKMFMSLFWVIGACIVYLLATPIAIIIPPLGLIYFFVORFVYASSROL 1140
DB 1081 DTVDSMIPETVYKMFMSLFWVIGACIVYLLATPIAIIIPPLGLIYFFVORFVYASSROL 1140
QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCTVLFALFAVSRHSLSAGLVGVSYSLOVTTYLNLVVRSSSEMTNIVA 1260
DB 1201 VRLCEVGNCTVLFALFAVSRHSLSAGLVGVSYSLOVTTYLNLVVRSSSEMTNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNVCYRREDLDFVLRHINTVINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNVCYRREDLDFVLRHINTVINGG 1320

QY 1321 EKVGIVRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLAREKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLAREKTIIPQDPVLF 1380
QY 1381 SSGIRNMLDPPSOYSDERWMTSLBLAKDFVSALPDKLDHECAEGENLSVGORQVYCL 1440
DB 1381 SSGIRNMLDPPSOYSDERWMTSLBLAKDFVSALPDKLDHECAEGENLSVGORQVYCL 1440
QY 1441 ARALIRKTIIVLEBATAVLETDLDLQSTIRIQEEDCTVLTIAHRLNTIMDYTRVYL 1500
DB 1441 ARALIRKTIIVLEBATAVLETDLDLQSTIRIQEEDCTVLTIAHRLNTIMDYTRVYL 1500
QY 1501 DKGEIOEYGAUSDLLQGRGLFYMAKADAGLV 1531
DB 1501 DKGEIOEYGAUSDLLQGRGLFYMAKADAGLV 1531

RESULT 7
US-08-141-893-2
Sequence 2, Application US/08141893
Patent No. 5489519
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deele, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/141,893
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923; 08/029,340
FILING DATE: 27-OCT-1992; 8-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-141-893-2

Query Match 80.6%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCAGSDGDPIDMNVNTWNTSPDFTKCFONTVLVWVCPYLIMACFPFYLYLSRH 60
DB 1 MALRGFCAGSDGSDPIMDNVNTWNTSNPDFTKCFONTVLVWVCPYLIMACFPFYLYLSRH 60
QY 61 DRGYQMPILNKTALGELTLMIVCMADLFYSFWERSRGI ELAPVIVSPITLIGITTLA 120
DB 61 DRGYQMPILNKTALGELTLMIVCMADLFYSFWERSRGI ELAPVIVSPITLIGITTLA 120

QY 121 TFLIQERRRGVSSGIMLTFMWALVCAALILSRKMTALKEBAQVDLPEDITFYFFS 180
DB 121 TFLIQERRRGVSSGIMLTFMWALVCAALILSRKMTALKEBAQVDLPEDITFYFFS 180
QY 181 LLLIQVLSCPSDSSPLFSETIHDNCPSSSASFSLRIFPMWTGLIVGYRQPLGSD 240
DB 181 LLLIQVLSCPSDSSPLFSETIHDNCPSSSASFSLRIFPMWTGLIVGYRQPLGSD 240
QY 241 LMSLNKEDTSQVVPVIVKWKKECAKTRKQPVVYVSSKDPAPOKESSKVDAEEVAL 300
DB 241 LMSLNKEDTSQVVPVIVKWKKECAKTRKQPVVYVSSKDPAPOKESSKVDAEEVAL 300
QY 301 IVKSPQKEMNPSLFKVLTKTGPYFLMSFPFKAIHDLMPSPQILKLLIKFVNDTQAPD 360
DB 301 IVKSPQKEMNPSLFKVLTKTGPYFLMSFPFKAIHDLMPSPQILKLLIKFVNDTQAPD 360
QY 361 WQGYFYTVLLFVTAQCLQTLVHOYFHIQFVSGMKIKTRAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGYFYTVLLFVTAQCLQTLVHOYFHIQFVSGMKIKTRAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLMNLGPSVLGAVAVWVLMVNVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLMNLGPSVLGAVAVWVLMVNVN 480
QY 481 AVMAKTKTYQVAMKSKDNRIKLMETLNGIKVLYAMELAKDKYLAIRQELKYLK 540
DB 481 AVMAKTKTYQVAMKSKDNRIKLMETLNGIKVLYAMELAKDKYLAIRQELKYLK 540
QY 541 KSAVLSAVGFTWCTPPLVALCTFAVYVITDENNIIDAQTAFSLALFNLRPLNLTLP 600
DB 541 KSAVLSAVGFTWCTPPLVALCTFAVYVITDENNIIDAQTAFSLALFNLRPLNLTLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVNAATFWARSDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVNAATFWARSDPT 660
QY 661 LMGITFSIPBGALVAVVQVCGKSLSLALAEKDEKGVHAIKGSVAVYPOQAMIOND 720
DB 661 LMGITFSIPBGALVAVVQVCGKSLSLALAEKDEKGVHAIKGSVAVYPOQAMIOND 720
QY 721 SLRENILPGCOLERPYSRVQAQALLPDLIELSGDRTLEGKGVNLSSGOKRVSILAR 780
DB 721 SLRENILPGCOLERPYSRVQAQALLPDLIELSGDRTLEGKGVNLSSGOKRVSILAR 780
QY 781 AVYSNADIVLEDDPLSAVDAVHGKIFENYIGPKGMLKNKRIIVTSHMSYLPQVDVIV 840
DB 781 AVYSNADIVLEDDPLSAVDAVHGKIFENYIGPKGMLKNKRIIVTSHMSYLPQVDVIV 840
QY 841 MSGKISEMSGYOELIARDGAFAPFLRTVASTEOBDAEENGVTGVSQPKAKOMENG 900
DB 841 MSGKISEMSGYOELIARDGAFAPFLRTVASTEOBDAEENGVTGVSQPKAKOMENG 900
QY 901 LVTPSAGKOLOROISSSSYSGGDISRHNSITAELOKAEKKEETWKMEADKAQGVYCL 960
DB 901 LVTPSAGKOLOROISSSSYSGGDISRHNSITAELOKAEKKEETWKMEADKAQGVYCL 960
QY 961 SVYMDYMAIGLFTSIFLFLMCHNVSALASNYWLSMTDPIYNGQETHKRYLSYVG 1020
DB 961 SVYMDYMAIGLFTSIFLFLMCHNVSALASNYWLSMTDPIYNGQETHKRYLSYVG 1020
QY 1021 ALGISQGIADVGYSAVSIIGIILASRCLHVDLHLSILSPSPFERTPSGMLVNRFSKEL 1080
DB 1021 ALGISQGIADVGYSAVSIIGIILASRCLHVDLHLSILSPSPFERTPSGMLVNRFSKEL 1080
QY 1081 DTVDSMIPEVIMKMGSLFNVIAGACIVYLLATPILAIIPPLGLIFPVQRFYVASSQOL 1140
DB 1081 DTVDSMIPEVIMKMGSLFNVIAGACIVYLLATPILAIIPPLGLIFPVQRFYVASSQOL 1140
QY 1141 KRLBSVRSPPYSHNETLIGSVIRAFEBQERFIHOSDLKYDENOKAYYSIVANRMLA 1200
DB 1141 KRLBSVRSPPYSHNETLIGSVIRAFEBQERFIHOSDLKYDENOKAYYSIVANRMLA 1200
QY 1201 VRLCEVCNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVVRMSEMETNIVA 1260

|||||
Db 1201 VRLCEVGNCLVLFALFAVIRSRHSLSAGLVGSLVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
|||
Qy 1261 VERLKEVSETEKEAPWQIQETAPRPSNPQVGRVERRNCIARREDLVLRHINVTNGG 1320
|||
Db 1261 VERLKEVSETEKEAPWQIQETRPPSPQVORVERRNCIARREDLVLRHINVTNGG 1320
|||
Qy 1321 EKVGIVGTGAGKSLTLGLFRINSEAGEIIIDGINIAKIGLHDLRFKTIITIPQDPVLF 1380
|||
Db 1321 EKVGIVGTGAGKSLTLGLFRINSEAGEIIIDGINIAKIGLHDLRFKTIITIPQDPVLF 1380
|||
Qy 1381 SGSLRMLNDPFSQYSDSEVWTSLELAHLKDFVSALPDCLDHECAGEGEMLSVGQROLVCL 1440
|||
Db 1381 SGSLRMLNDPFSQYSDSEVWTSLELAHLKDFVSALPDCLDHECAGEGEMLSVGQROLVCL 1440
|||
Qy 1441 ARALLRKTIVLDEATAVDLETDLIQSTIRTOFEDCTVLTAAHRIINTMDYTRIVL 1500
|||
Db 1441 ARALLRKTIVLDEATAVDLETDLIQSTIRTOFEDCTVLTAAHRIINTMDYTRIVL 1500
|||
Qy 1501 DKGEIOEYGAAPSDLLQORGLFYMAKQAGLV 1531
|||
Db 1501 DKGEIOEYGAAPSDLLQORGLFYMAKQAGLV 1531
|||

RESULT 8
US-08-463-0928-2
; Sequence 2, Application US/084630928
; Patent No. 5766880
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P. C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,0928
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-0928-2

Query Match 80.6%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDANTANTSNDFPKCFONTLVWVPCFLMACFPFYLYSRH 60
|||
Db 1 MALRGFCSADGSDPLMDANTANTSNDFPKCFONTLVWVPCFLMACFPFYLYSRH 60
|||
Qy 61 DRGYIQMTPLNKTALGFLMIYVWADLFYSFWERSRGIFLAPVFLVPTLGIITTLA 120
|||
Db 61 DRGYIQMTPLNKTALGFLMIYVWADLFYSFWERSRGIFLAPVFLVPTLGIITTLA 120
|||
Qy 121 TFLIQLERRKGVQSSGIMLTFWLVYCALAIIKSKIMTLAKEDAQVDLFRDITFYVFS 180
|||
Db 121 TFLIQLERRKGVQSSGIMLTFWLVYCALAIIKSKIMTLAKEDAQVDLFRDITFYVFS 180
|||
Qy 181 LLIILQVLSGFSRSPLEFETIHDPNCPBSSASFISRIITFWMTGLIVRGYRPLGSD 240
|||
Db 181 LLIILQVLSGFSRSPLEFETIHDPNCPBSSASFISRIITFWMTGLIVRGYRPLGSD 240
|||
Qy 241 LMSLNKEDTSEOVVPLVKNWKKCECAKTRKQPVKVVSSKDPAQPKESKVDANEVEAL 300
|||
Db 241 LMSLNKEDTSEOVVPLVKNWKKCECAKTRKQPVKVVSSKDPAQPKESKVDANEVEAL 300
|||
Qy 301 IVKSPQKEMNPISLFKVLKTFGEYFLMSFPFKAIHDIIMFSGPQILKLIKPVNDYAPD 360
|||
Db 301 IVKSPQKEMNPISLFKVLKTFGEYFLMSFPFKAIHDIIMFSGPQILKLIKPVNDYAPD 360
|||
Qy 361 WQGYFYTVLLFVTAQOTLVHQQYFHTCFVSGRIKTAIVGAYRKALVTNSRKSTV 420
|||
Db 361 WQGYFYTVLLFVTAQOTLVHQQYFHTCFVSGRIKTAIVGAYRKALVTNSRKSTV 420
|||
Qy 421 GEIVNLMSVDQRFMDLATYINMIWSAPLOVIAIYLMNLGSPVLAGVAVMWLVPN 480
|||
Db 421 GEIVNLMSVDQRFMDLATYINMIWSAPLOVIAIYLMNLGSPVLAGVAVMWLVPN 480
|||
Qy 481 AVNAAKTKTYQVAHMSKDNRIKLANEILNGIVLKLIVANELAKDKVLAIROBELKVLK 540
|||
Db 481 AVNAAKTKTYQVAHMSKDNRIKLANEILNGIVLKLIVANELAKDKVLAIROBELKVLK 540
|||
Qy 541 KSAVYLSAVGTFTWVCTPVLALCTPAVYVITIDNNIILDAQTAVSLAFNIIIRPPLNLP 600
|||
Db 541 KSAVYLSAVGTFTWVCTPVLALCTPAVYVITIDNNIILDAQTAVSLAFNIIIRPPLNLP 600
|||
Qy 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVDDGGTNSITVRNAFTTARSDPT 660
|||
Db 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVDDGGTNSITVRNAFTTARSDPT 660
|||
Qy 661 LNCITTSIPBGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVVPOQAIQND 720
|||
Db 661 LNCITTSIPBGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVVPOQAIQND 720
|||
Qy 721 SLAENILFGQLEBPYRSTYQACALLPDLIELPSGDRTEIGKGVNLSSGQQRVSILAR 780
|||
Db 721 SLAENILFGQLEBPYRSTYQACALLPDLIELPSGDRTEIGKGVNLSSGQQRVSILAR 780
|||
Qy 781 AVYSNADIVLPDPLSAVDARVGHKIFENVIGKGMKNKTRILVTHSMYSILPOVDYIV 840
|||
Db 781 AVYSNADIVLPDPLSAVDARVGHKIFENVIGKGMKNKTRILVTHSMYSILPOVDYIV 840
|||
Qy 841 MSGGKISEMGSYQELLARDGAFELFTYASTQEOBDAENGYTGVSGPKKAKOMENGM 900
|||
Db 841 MSGGKISEMGSYQELLARDGAFELFTYASTQEOBDAENGYTGVSGPKKAKOMENGM 900
|||
Qy 901 LVITDSAGKQQLQRLSSSSVSGDISRHHSNTAEILOAEAKKEFTWKLMEADKQGTGVKL 960
|||

Db 901 LVTSAGQLOKROSLSSSSSSSGDLSRHNSIAELQKAKKEETWKMEAKQGYKL 960
Qy 961 SVYDYMKAIGLFTISFLIFMCMHVSALASNYLSLMTDDPIYNGQHTKRLSYG 1020
Db 961 SVYDYMKAIGLFTISFLIFMCMHVSALASNYLSLMTDDPIYNGQHTKRLSYG 1020
Qy 1021 ALGISQGIAGVGYMANVSGIILASRCLHVDLHLSILSPMSPEERTPSGNLVNRSKEL 1080
Db 1021 ALGISQGIAGVGYMANVSGIILASRCLHVDLHLSILSPMSPEERTPSGNLVNRSKEL 1080
Qy 1081 DTUSMTPEVTKMGSLFNVIAGACIVILLATPIAIIIPGLIYFVQGFYVASSQOL 1140
Db 1081 DTUSMTPEVTKMGSLFNVIAGACIVILLATPIAIIIPGLIYFVQGFYVASSQOL 1140
Qy 1141 KRLSVSRSPYSHFNETLGVSVIRAFEOERFIHQSDLVKVDENQKAYPSIVANRMLA 1200
Db 1141 KRLSVSRSPYSHFNETLGVSVIRAFEOERFIHQSDLVKVDENQKAYPSIVANRMLA 1200
Qy 1201 VRLCEVGCIVLPALPAVISRHSISAGLVGLSVSLQVTTYINMLVVRMSSEMETNIVA 1260
Db 1201 VRLCEVGCIVLPALPAVISRHSISAGLVGLSVSLQVTTYINMLVVRMSSEMETNIVA 1260
Qy 1261 VERKEVEETEKEAPMOIOETAPPSWPQVGRVEFRNYCLRYREDLDFVLHINVTNGG 1320
Db 1261 VERKEVEETEKEAPMOIOETAPPSWPQVGRVEFRNYCLRYREDLDFVLHINVTNGG 1320
Qy 1321 EKVGI VGTGAKGSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPDPPLVF 1380
Db 1321 EKVGI VGTGAKGSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPDPPLVF 1380
Qy 1381 SGLSRMNDLPPSQYSDDEEWTSLBLAHLKDPVSALPDKLDHCAEGENLSVGQRLVCL 1440
Db 1381 SGLSRMNDLPPSQYSDDEEWTSLBLAHLKDPVSALPDKLDHCAEGENLSVGQRLVCL 1440
Qy 1441 ARALRKTKIIVLBATAVNDLETDLIQSTIRTPROFECTYLTANRANTIMDYRIVVL 1500
Db 1441 ARALRKTKIIVLBATAVNDLETDLIQSTIRTPROFECTYLTANRANTIMDYRIVVL 1500
Qy 1501 DKGEIOEYGAPSDLLQORGLFFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQORGLFFYSMAKDAGLV 1531

RESULT 9

US-08-462-109A-2

Sequence 2, Application US/08462109A

Patent No. 5882875

GENERAL INFORMATION:

APPLICANT: Cole, Susan P.C.

APPLICANT: Dealey, Roger G.

TITLE OF INVENTION: METHODS FOR IDENTIFYING

TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: LAHIVE & COCKFIELD

STREET: 60 State Street, suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08462,109A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/966,923

FILING DATE: 27-OCT-1992

APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP4
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-2

Query Match 80.6%; Score 7849; DB 1; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGCSADGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCFYLMACFPFYLYLSRH 60
Db 1 MALRGCSADGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCFYLMACFPFYLYLSRH 60
Qy 61 DRGIQNTPLNKTALGFLIMTYCMAFLFYSFWERSRGIFLAVFLVSPPLGLITLLA 120
Db 61 DRGIQNTPLNKTALGFLIMTYCMAFLFYSFWERSRGIFLAVFLVSPPLGLITLLA 120
Qy 121 TFLQLRRKGVQSSGIMLTFWLVALCALILRSKIMTALKEBDAVDLFRDITFYVYS 180
Db 121 TFLQLRRKGVQSSGIMLTFWLVALCALILRSKIMTALKEBDAVDLFRDITFYVYS 180
Qy 181 LLLIQVLVSCFSDRSPFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCFSDRSPFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240
Qy 241 LMSLNKEDTSQVVPVLVKNKKKCACTRKQPVVYVSSSDPAOPKSSSKVDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVLVKNKKKCACTRKQPVVYVSSSDPAOPKSSSKVDANEVEAL 300
Qy 301 IVKSPKEMNPSLKVLYKTPGFYFLMSFFPKAIDHLMBSGPQILKLIKFNVDTKAPD 360
Db 301 IVKSPKEMNPSLKVLYKTPGFYFLMSFFPKAIDHLMBSGPQILKLIKFNVDTKAPD 360
Qy 361 WQGYFTVLLFVTAQLQTLVLAQYFHIQVSGMRKIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTAQLQTLVLAQYFHIQVSGMRKIKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAQRMIDLATTYINMTWSAPLOYITALYILMLNLSGSVLAGAVMYLAMPVN 480
Db 421 GEIYNLMSVDAQRMIDLATTYINMTWSAPLOYITALYILMLNLSGSVLAGAVMYLAMPVN 480
Qy 481 AYMAKTKTYOVAMHKSNDRIKLMNRLINGIKYLKLYAMELAKDKVYLAIRBELKYLK 540
Db 481 AYMAKTKTYOVAMHKSNDRIKLMNRLINGIKYLKLYAMELAKDKVYLAIRBELKYLK 540
Qy 541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNIIDAQTAFFSALFNILRFPPLNLP 600
Db 541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNIIDAQTAFFSALFNILRFPPLNLP 600
Qy 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGTNSITVENATFTMARSDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGTNSITVENATFTMARSDPT 660
Qy 661 LNGITFSTPEGALVAVVGQVCGKSSILSLAAMDVEGHVAITKSAVAYVPOQAWIOND 720
Db 661 LNGITFSTPEGALVAVVGQVCGKSSILSLAAMDVEGHVAITKSAVAYVPOQAWIOND 720

QY 721 SLRNNILFGCOLLEPPYRSVIOACALLPDLLEILPSGDRTEIGEKNVLSGGOKORVSLAR 780
| | | | |
Db 721 SLRNNILFGCOLLEPPYRSVIOACALLPDLLEILPSGDRTEIGEKNVLSGGOKORVSLAR 780
QY 781 AVYSNADIIYLFDDPLSAVDAAVGHKIFENVIQPKMLKNKTRILLVTHSMYSLPOVDVILV 840
| | | | |
Db 781 AVYSNADIIYLFDDPLSAVDAAVGHKIFENVIQPKMLKNKTRILLVTHSMYSLPOVDVILV 840
QY 841 MSGGKISEMGSYOELLARDGAFAPFLRTYASTEOBDAEENGVCVSGPGKEAQMENG 900
| | | | |
Db 841 MSGGKISEMGSYOELLARDGAFAPFLRTYASTEOBDAEENGVCVSGPGKEAQMENG 900
QY 901 LVTDSAGKOLOROSSSSSSSGDISRHHNSTAELOKAEKEEPMKMEADKAOTGOVKL 960
| | | | |
Db 901 LVTDSAGKOLOROSSSSSSSGDISRHHNSTAELOKAEKEEPMKMEADKAOTGOVKL 960
QY 961 SVYWDYKKAIGLFTISFLSIFLMCNHVSALASNYWLSMTDPIVNGTOEHTKYRLSVYG 1020
| | | | |
Db 961 SVYWDYKKAIGLFTISFLSIFLMCNHVSALASNYWLSMTDPIVNGTOEHTKYRLSVYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLRSRSPMSFPERTPSGULVNRFSKEL 1080
| | | | |
Db 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLRSRSPMSFPERTPSGULVNRFSKEL 1080
QY 1081 DTVDMSIPEVIKMGSLFNVIAGACIVILATPIAIIIPGLIYFVORFVYASSROL 1140
| | | | |
Db 1081 DTVDMSIPEVIKMGSLFNVIAGACIVILATPIAIIIPGLIYFVORFVYASSROL 1140
QY 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANWMLA 1200
| | | | |
Db 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANWMLA 1200
QY 1201 VRELCVNCIVLPALPAVIRSRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
| | | | |
Db 1201 VRELCVNCIVLPALPAVIRSRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPPSMPQVGRVFRNYCCLRYREDLPLVLRHINVTNGG 1320
| | | | |
Db 1261 VERLKEYSETEKEAPMOIOETAPPSMPQVGRVFRNYCCLRYREDLPLVLRHINVTNGG 1320
QY 1321 EKVGIIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIITIPQDPVLF 1380
| | | | |
Db 1321 EKVGIIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIITIPQDPVLF 1380
QY 1381 SGRSRMNLDPSSQYSDSEWTSLELAHLKDFVSLPDKLDHECAGEGENISVQOROLVCL 1440
| | | | |
Db 1381 SGRSRMNLDPSSQYSDSEWTSLELAHLKDFVSLPDKLDHECAGEGENISVQOROLVCL 1440
QY 1441 ABALLRKTJILVDEATPAVDLEETDILQSTIRFOFEDCTVLTIAHRLNTIMOTRYIVL 1500
| | | | |
Db 1441 ABALLRKTJILVDEATPAVDLEETDILQSTIRFOFEDCTVLTIAHRLNTIMOTRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531
| | | | |
Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531

RESULT 10
US-08-460-907B-2
; Sequence 2, Application US/08460907B
; Patent No. 5891724

; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P.C.
; TITLE OF INVENTION: METHODS FOR CONFERRING MDT/IDRUG
; TITLE OF INVENTION: RESISTANCE ON A CELL
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTBO RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA

ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-2

Query Match 80.6%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVNTSNPDFTKCFQNTLVVWVPCFYLMACEPFYFLYSRH 60
| | | | |
Db 1 MALRGFCSADGSDPLMDMNTVNTSNPDFTKCFQNTLVVWVPCFYLMACEPFYFLYSRH 60
QY 61 DRGVIQMTPLNKTCTALGFLIMIVCWADLFYSFWERSRGIFLAPVFLVSPLLGITTLLA 120
| | | | |
Db 61 DRGVIQMTPLNKTCTALGFLIMIVCWADLFYSFWERSRGIFLAPVFLVSPLLGITTLLA 120
QY 121 TPLIOLERRKGVOSGGLMFLFVVALVCAIIRSKIMTALKXDAQVDLFRDITFYVYFS 180
| | | | |
Db 121 TPLIOLERRKGVOSGGLMFLFVVALVCAIIRSKIMTALKXDAQVDLFRDITFYVYFS 180
QY 181 LLLIOLVLSGFSRSPPLFSETIHDPNPCPESSASFLSLTFMWTGILVAGYRQPLEGSD 240
| | | | |
Db 181 LLLIOLVLSGFSRSPPLFSETIHDPNPCPESSASFLSLTFMWTGILVAGYRQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAPQKSSKVDANEEVEAL 300
| | | | |
Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAPQKSSKVDANEEVEAL 300
QY 301 IVKSPQKSNPSSLFKVLYKTPGPFYFLMSPFKAIHDLMPGSGQIILKLIKFNDRKAPD 360
| | | | |
Db 301 IVKSPQKSNPSSLFKVLYKTPGPFYFLMSPFKAIHDLMPGSGQIILKLIKFNDRKAPD 360
QY 361 WQGYFTVLLFVTAQOTLVLAHQYFHCIVSGNRIRKTAIVIGAYRRALVITNSARKSSTV 420
| | | | |
Db 361 WQGYFTVLLFVTAQOTLVLAHQYFHCIVSGNRIRKTAIVIGAYRRALVITNSARKSSTV 420

QY 421 GEIVNLSVDAQRFMDLATYINMISAPLOYIALYLLMLNIGPSVLAVAVMVLMPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMISAPLOYIALYLLMLNIGPSVLAVAVMVLMPVN 480
QY 481 AVAMAKTKTYOVAMHKSNDRIKLMNELNGLIKYLKYLAWELAPFOKYLAROEELKYLK 540
DB 481 AVAMAKTKTYOVAMHKSNDRIKLMNELNGLIKYLKYLAWELAPFOKYLAROEELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNIIDPAOTAFVSLFENILFPLNITLP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNIIDPAOTAFVSLFENILFPLNITLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
QY 661 LNGTHFSPREGALVAVVQVCGKSLLSALLAENDKVEGHVALKGSVAVYPOQAMIOND 720
DB 661 LNGTHFSPREGALVAVVQVCGKSLLSALLAENDKVEGHVALKGSVAVYPOQAMIOND 720
QY 721 SLRENILFCOLEEPPYRSVIOACALPDLLEILPSGDRTEIGEKNVLSGQKORVSLAR 780
DB 721 SLRENILFCOLEEPPYRSVIOACALPDLLEILPSGDRTEIGEKNVLSGQKORVSLAR 780
QY 781 AVYSNADIIYFLDDPLSAVDAHVKHIFENVIGPKMLNKXTRILVTHSMGYLPQVDVII 840
DB 781 AVYSNADIIYFLDDPLSAVDAHVKHIFENVIGPKMLNKXTRILVTHSMGYLPQVDVII 840
QY 841 MSGKISMSGYSVOELLARDGAFAEFLRYASTEOQDAEENGVCVSGPGEAKEMGCM 900
DB 841 MSGKISMSGYSVOELLARDGAFAEFLRYASTEOQDAEENGVCVSGPGEAKEMGCM 900
QY 901 LVTSAGKOLOROLSSSSSSSGDISRHNSSTAELOKAEKKEETWKLMEADKAOTQYKL 960
DB 901 LVTSAGKOLOROLSSSSSSSGDISRHNSSTAELOKAEKKEETWKLMEADKAOTQYKL 960
QY 961 SVYWDYMAIGLIFISFISIFLPMCNCVSAASNYMLSLMTDDPIVNGOETHKVALSVYG 1020
DB 961 SVYWDYMAIGLIFISFISIFLPMCNCVSAASNYMLSLMTDDPIVNGOETHKVALSVYG 1020
QY 1021 ALGISQGIAPFGYSNAVSIIGGILARCLHDLHLSILSPMSFFERTSGNUNRFSKEL 1080
DB 1021 ALGISQGIAPFGYSNAVSIIGGILARCLHDLHLSILSPMSFFERTSGNUNRFSKEL 1080
QY 1081 DTVDSMIEVIMKMFMSLFPNVIGACIVILATPIAIIIPPLGLIYFFVQRPYVASSROL 1140
DB 1081 DTVDSMIEVIMKMFMSLFPNVIGACIVILATPIAIIIPPLGLIYFFVQRPYVASSROL 1140
QY 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEORFIHQSDLKVDENOKAYYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEORFIHQSDLKVDENOKAYYPSIVANRWLA 1200
QY 1201 VRLEVCVGCIVLFAALFAVISRHSLSAGIVGVSYSIQVTTYLWMLVRMSSEMTINIVA 1260
DB 1201 VRLEVCVGCIVLFAALFAVISRHSLSAGIVGVSYSIQVTTYLWMLVRMSSEMTINIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETRPSPSPQVGRVERRNCLARKREDLDFVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPMOIOETRPSPSPQVGRVERRNCLARKREDLDFVLRHINVTINGG 1320
QY 1321 EKVGVIGRTGAGKSSLTGLFPRINESAGEIIIDGINIAKIGLHDLRFKITIIPDPVYLF 1380
DB 1321 EKVGVIGRTGAGKSSLTGLFPRINESAGEIIIDGINIAKIGLHDLRFKITIIPDPVYLF 1380
QY 1381 SCSLRPMNDPFSQYSDEBEVMTSLBLAHLKDFVSALPDKLDHECAGGENLSVGQOLVCL 1440
DB 1381 SCSLRPMNDPFSQYSDEBEVMTSLBLAHLKDFVSALPDKLDHECAGGENLSVGQOLVCL 1440
QY 1441 ABALIRKTKIIVLADATAAVDLETDLIQSTIRTOFECTVLTIAHRLNTIMDYRIVYL 1500
DB 1441 ABALIRKTKIIVLADATAAVDLETDLIQSTIRTOFECTVLTIAHRLNTIMDYRIVYL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
RESULT 11
US-08-463-179A-2
Sequence 2, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-2
Query Match 80.6%; Score 7849; DB 2; Length 1531;
Best local similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALGFCSADGSDLLMDMNVMTWNTSNPDFTKCFONTYLVWPCFYLVACFPFELYLSRH 60
DB 1 MALGFCSADGSDLLMDMNVMTWNTSNPDFTKCFONTYLVWPCFYLVACFPFELYLSRH 60
QY 61 DRGYIOMTPLNKTATGALGFLIMIVCMADLFYSFMRSGIFLAVFLVSPILLGITTLA 120
DB 61 DRGYIOMTPLNKTATGALGFLIMIVCMADLFYSFMRSGIFLAVFLVSPILLGITTLA 120
QY 121 TFLIQLERKKGVSQSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180
DB 121 TFLIQLERKKGVSQSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180
QY 181 LLLIQLVLSGFSDDSPFLSETIHDNPPCPSSASFLSITTWMTTGLIVRGYROPLESD 240
DB 181 LLLIQLVLSGFSDDSPFLSETIHDNPPCPSSASFLSITTWMTTGLIVRGYROPLESD 240

QY 241 LMSLNKEDTSEQVVPVLVKNKKKCEAKTRKQPVKVYSSKDPAPQKSSKYDANEAEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNKKKCEAKTRKQPVKVYSSKDPAPQKSSKYDANEAEAL 300
QY 301 IVKSPQKEMNPSLFEVLYKTFGPYFLMSFFPKAIDHLMFSGPOLIKLIFVNDYKAD 360
DB 301 IVKSPQKEMNPSLFEVLYKTFGPYFLMSFFPKAIDHLMFSGPOLIKLIFVNDYKAD 360
QY 361 WQGFYVTLFVTACLOTLVHOYFHCIVSGMRLKTAIVGAVYRKALVITNSARKSTV 420
DB 361 WQGFYVTLFVTACLOTLVHOYFHCIVSGMRLKTAIVGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDQRFMDLATYINMISAPLOVILALYLLMLNGPSVLGAVVWMLMPVN 480
DB 421 GEIVNLMSVDQRFMDLATYINMISAPLOVILALYLLMLNGPSVLGAVVWMLMPVN 480
QY 481 AVMAKTKTYOAHMKSNDRIKLMNEILNGIKVYLKAYMELAFKDKVLAIROBELKYLK 540
DB 481 AVMAKTKTYOAHMKSNDRIKLMNEILNGIKVYLKAYMELAFKDKVLAIROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVYIDENNILDAQTAFLALFNILRPLNLP 600
DB 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVYIDENNILDAQTAFLALFNILRPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGCTNSITVRNATFTMARSDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGCTNSITVRNATFTMARSDPT 660
QY 661 LMGITFSPPEGALVAVVQVCGKSSLSLALLAEMDKVEGHVAKGSVAVYPOQAWIOND 720
DB 661 LMGITFSPPEGALVAVVQVCGKSSLSLALLAEMDKVEGHVAKGSVAVYPOQAWIOND 720
QY 721 SLRNNIIFGCOLLEPPYRSVYQACALLPDLLEILPSGRTEIGEKVNLSSGQKRVSLAR 780
DB 721 SLRNNIIFGCOLLEPPYRSVYQACALLPDLLEILPSGRTEIGEKVNLSSGQKRVSLAR 780
QY 781 AVVSNADITYLPDDLSAVDAHVGHIFENVIGPKMLKNKRIIVTMSMSTLPQVDYIV 840
DB 781 AVVSNADITYLPDDLSAVDAHVGHIFENVIGPKMLKNKRIIVTMSMSTLPQVDYIV 840
QY 841 MSGGKISEMGSYQELLRDGAFAEFLRTYASTEODAEENGYVSGSPGEAKQOMENG 900
DB 841 MSGGKISEMGSYQELLRDGAFAEFLRTYASTEODAEENGYVSGSPGEAKQOMENG 900
QY 901 LVTDSAGKOLOROLSSSSSSYSGDISRHNSTAELOKAEKKEETWKLMEADKAQTGYKL 960
DB 901 LVTDSAGKOLOROLSSSSSSYSGDISRHNSTAELOKAEKKEETWKLMEADKAQTGYKL 960
QY 961 SYVVDYMKATGLFISFLSTPLFMCNHVSALASNTWLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SYVVDYMKATGLFISFLSTPLFMCNHVSALASNTWLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVGYMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGULVNFSEKL 1080
DB 1021 ALGISOGIAVGYMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGULVNFSEKL 1080
QY 1081 DTVDSMIPEVILKMFMSLFNVIAGACIVILATPIAIIIPGLIYEFVQRFYVASSROL 1140
DB 1081 DTVDSMIPEVILKMFMSLFNVIAGACIVILATPIAIIIPGLIYEFVQRFYVASSROL 1140
QY 1141 KRLESVRSPPYSHFNFTLGVSTIRAFEEBERIHOSDLKVDENOKAYPSIYANWLA 1200
DB 1141 KRLESVRSPPYSHFNFTLGVSTIRAFEEBERIHOSDLKVDENOKAYPSIYANWLA 1200
QY 1201 VRLCEVNCIYLFALFAVISRHSLSAGLVGLSVSYLQVTTYANMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLFALFAVISRHSLSAGLVGLSVSYLQVTTYANMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPKMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHINVTING 1320
DB 1261 VERLKEYSETEKEAPKMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHINVTING 1320
QY 1321 EKVGIVERTGAGKSSLTGLFRINESAGEIIDIIGINIAKGLHDLAFKTIIPQDVL 1380

DB 1321 EKVGIVERTGAGKSSLTGLFRINESAGEIIDIIGINIAKGLHDLAFKTIIPQDVL 1380
QY 1381 SGSLRNMLDPPSOYSDDEEVTSLBLAHLKQFVSALPKLDHECAGEGENTSVGQROLVCL 1440
DB 1381 SGSLRNMLDPPSOYSDDEEVTSLBLAHLKQFVSALPKLDHECAGEGENTSVGQROLVCL 1440
QY 1441 ABALRRTKILVDEATAVADLETDLIQSTIRQFEDCTVLTARHLNTIMDYTRIVL 1500
DB 1441 ABALRRTKILVDEATAVADLETDLIQSTIRQFEDCTVLTARHLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDILQORGLFYSMAKDAGLV 1531
RESULT 12
US-08-461-384B-2
Sequence 2, Application US/08461384B
Patent No. 6025473
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,384B
FILING DATE: 05-JUN-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01547
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEO ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-384B-2
Query Match 80.64; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.94; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTWNTSNPDTKCFQNTVLVWPCFYLMACPPFYFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNTWNTSNPDTKCFQNTVLVWPCFYLMACPPFYFLYLSRH 60
QY 61 DRGIVQMTPLNKTKTALGFLMIVCANADLFYSEWERSRGIFLAPVFLVSTLLGITTILA 120

```

Db      61 DRGIQWTPLNKRTALGFLMIVCWLDFYSFWERSGIFLAPFLVSPFLITLLA 120
Qy      121 TFLQLERRKVOSSGIMLTFMVALVCALALIRSKITALKEDAOVDLFDITFYVFS 180
Db      121 TFLQLERRKVOSSGIMLTFMVALVCALALIRSKITALKEDAOVDLFDITFYVFS 180
Qy      181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRTFMWITGLIVRGYROPLEGSD 240
Db      181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRTFMWITGLIVRGYROPLEGSD 240
Qy      241 LMSLNKEDTSEQVVPVLYKMKKGCATRKOPKVKVYSSKDPAPQKSSKVDANEVVAL 300
Db      241 LMSLNKEDTSEQVVPVLYKMKKGCATRKOPKVKVYSSKDPAPQKSSKVDANEVVAL 300
Qy      301 IVKSPQKEMNPSLRFVLYKTFGPFYLMSEFPFAIHDLMFSGPOLIKLLIFVNDTKAPD 360
Db      301 IVKSPQKEMNPSLRFVLYKTFGPFYLMSEFPFAIHDLMFSGPOLIKLLIFVNDTKAPD 360
Qy      361 WQGYFYTVLFPVTAQLQTLVLAHQYFHCIVSGMRIKTAVIGA VYRKALVITNSARKSSTV 420
Db      361 WQGYFYTVLFPVTAQLQTLVLAHQYFHCIVSGMRIKTAVIGA VYRKALVITNSARKSSTV 420
Qy      421 GEIYNLMSVDAQRFMDLATYINMISAPLOVILALYLLMIMGSVLAGVAVMVLMPVFN 480
Db      421 GEIYNLMSVDAQRFMDLATYINMISAPLOVILALYLLMIMGSVLAGVAVMVLMPVFN 480
Qy      481 AVMAMKTKYOVAAHMKSKDNRIKLMNEILNGIKVLKYLAMELAFDKYLAIROEELKYLK 540
Db      481 AVMAMKTKYOVAAHMKSKDNRIKLMNEILNGIKVLKYLAMELAFDKYLAIROEELKYLK 540
Qy      541 KSAVLSAVGTFETWCTPFLVALCTFAVYVITDENNILDQAFAVSLAFNLLRPENLILP 600
Db      541 KSAVLSAVGTFETWCTPFLVALCTFAVYVITDENNILDQAFAVSLAFNLLRPENLILP 600
Qy      601 MVISSIVASVSLKRLRIFLSHEELBPSIERRPKYOGGCTNSTVVRATFTWASDPT 660
Db      601 MVISSIVASVSLKRLRIFLSHEELBPSIERRPKYOGGCTNSTVVRATFTWASDPT 660
Qy      661 LINGITFSLPEGALVAVVGVGCGKSLLSALLAEMDKVGHAIKSAVYVPOQAMIOND 720
Db      661 LINGITFSLPEGALVAVVGVGCGKSLLSALLAEMDKVGHAIKSAVYVPOQAMIOND 720
Qy      721 SLRENILFGCOLBEEPYRSVIOACALPDLBILPSGDRTEIGEKVNLSGGOKORVSLAR 780
Db      721 SLRENILFGCOLBEEPYRSVIOACALPDLBILPSGDRTEIGEKVNLSGGOKORVSLAR 780
Qy      781 AVYSNADIYLPDDPLSAVDAAHGKHI FENVIGPKMLKXKTRILVTHSMSTLPQVDVILV 840
Db      781 AVYSNADIYLPDDPLSAVDAAHGKHI FENVIGPKMLKXKTRILVTHSMSTLPQVDVILV 840
Qy      841 MSGGKISMGSGVOELLARDGAFAEFLRTYASTEOQDABENGCVTVSGGKAKOMENGM 900
Db      841 MSGGKISMGSGVOELLARDGAFAEFLRTYASTEOQDABENGCVTVSGGKAKOMENGM 900
Qy      901 LVYDSAGKOROLSSSSSSSGDISRHHNSTALOKABAKKEETKMLEADYAKTGOYVL 960
Db      901 LVYDSAGKOROLSSSSSSSGDISRHHNSTALOKABAKKEETKMLEADYAKTGOYVL 960
Qy      961 SVYMDYMKAIGLFISFLSIFLPMCNHVSALASNYWLSMTDDPIVNGTQEHKVALSVYG 1020
Db      961 SVYMDYMKAIGLFISFLSIFLPMCNHVSALASNYWLSMTDDPIVNGTQEHKVALSVYG 1020
Qy      1021 ALGISQGIAPVGYMAVSIIGLILASRCLHVDLHLSILSPMSFERTSGNLYNFSKEL 1080
Db      1021 ALGISQGIAPVGYMAVSIIGLILASRCLHVDLHLSILSPMSFERTSGNLYNFSKEL 1080
Qy      1081 DTVDMSIPEVIMKFMGSLFENVIGACIVILLATPIAIIIPGLIYFFVQRFYVASSRQL 1140
Db      1081 DTVDMSIPEVIMKFMGSLFENVIGACIVILLATPIAIIIPGLIYFFVQRFYVASSRQL 1140
Qy      1141 KRLBSVSSSPVYSHNETLLGVSVIRAFEOERFIHOSDLKVDENOKAYPSIVANRWLA 1200

```

```

Db      1141 KRLBSVSSPVYSHNETLLGVSVIRAFEOERFIHOSDLKVDENOKAYPSIVANRWLA 1200
Qy      1201 VRLCEVNCNIVLFAALFAVISRHSLSAGLVGSYSLOVTTYLMLVMSSEMETNIVA 1260
Db      1201 VRLCEVNCNIVLFAALFAVISRHSLSAGLVGSYSLOVTTYLMLVMSSEMETNIVA 1260
Qy      1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFNRYCLRYRDLDPVLRHINVTINGG 1320
Db      1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFNRYCLRYRDLDPVLRHINVTINGG 1320
Qy      1321 EKVGIVERTAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Db      1321 EKVGIVERTAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Qy      1381 SGSLRNMLDPFSQYSDDEVWTSLELAHKDFVSAIPDKLDHECAGEGENLSVGGROLVCL 1440
Db      1381 SGSLRNMLDPFSQYSDDEVWTSLELAHKDFVSAIPDKLDHECAGEGENLSVGGROLVCL 1440
Qy      1441 ARALLRTKTLIVDEATPAVDLEFTDILLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
Db      1441 ARALLRTKTLIVDEATPAVDLEFTDILLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
Qy      1501 DKGEIOEYGAPSDILLQORGLFYSMAXDAGLV 1531
Db      1501 DKGEIOEYGAPSDILLQORGLFYSMAXDAGLV 1531

RESULT 13
US-08-407-207A-2
; Sequence 2, Application US/08407207A
; Patent No. 6063621
; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P.C.
; TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESS: PARTO RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,207A
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: 01512
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-407-207A-2

```


Query Match 80.6% Score 7849; DB 2; Length 1531;
 Best Local Similarity 99.9% Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALGFCGADSDPLMDNVTMTNTPKCFONTLVWVPCFYLMACPFYFYLSRH 60
 DB 1 MALGFCGADSDPLMDNVTMTNTPKCFONTLVWVPCFYLMACPFYFYLSRH 60
 QY 61 DRGTYQMTPLNKTALGFLMIYCMADLFSPWERSGIFLAPFVLSPTLLGTTLLA 120
 DB 61 DRGTYQMTPLNKTALGFLMIYCMADLFSPWERSGIFLAPFVLSPTLLGTTLLA 120
 QY 121 TELLQLERRKGVSSGIMLTFMLVALVICALILRSKITALKEDAOVDLPFDITFYVFS 180
 DB 121 TELLQLERRKGVSSGIMLTFMLVALVICALILRSKITALKEDAOVDLPFDITFYVFS 180
 QY 181 LLLIQLVLSGSDRSLPSETIHDNPPCESSASFLSRITFMWITGLIVRGYROPLEGSD 240
 DB 181 LLLIQLVLSGSDRSLPSETIHDNPPCESSASFLSRITFMWITGLIVRGYROPLEGSD 240
 QY 241 LMSLNKEDTSQOVVPLVKNKKKCAKTRKOPVXVYSSKDPAPKSSKYDANEYEAL 300
 DB 241 LMSLNKEDTSQOVVPLVKNKKKCAKTRKOPVXVYSSKDPAPKSSKYDANEYEAL 300
 QY 301 IVKSPOKEMNPSLFCULYKTEGYPFLMSFFKAIHDMFSGPOLIKLLIKFVNDTAPD 360
 DB 301 IVKSPOKEMNPSLFCULYKTEGYPFLMSFFKAIHDMFSGPOLIKLLIKFVNDTAPD 360
 QY 361 WQGYFYTVLFTVACLOTVLHQYFHLICFVSGMRIKTAIVAGYRKALVITNSARKSTV 420
 DB 361 WQGYFYTVLFTVACLOTVLHQYFHLICFVSGMRIKTAIVAGYRKALVITNSARKSTV 420
 QY 421 GEIYNLMSVDAQRFMDLATYINMTWSAPLOVILALYILMLNGSVLAGVAVMLAVPN 480
 DB 421 GEIYNLMSVDAQRFMDLATYINMTWSAPLOVILALYILMLNGSVLAGVAVMLAVPN 480
 QY 481 AVMMKTKTYOVAMKSKDNRIKLMNELINGIKYLKIYAMELARKDKYLAIRBELKYLK 540
 DB 481 AVMMKTKTYOVAMKSKDNRIKLMNELINGIKYLKIYAMELARKDKYLAIRBELKYLK 540
 QY 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVITDENNIIDAQAFVSLAFNIRPPLNLP 600
 DB 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVITDENNIIDAQAFVSLAFNIRPPLNLP 600
 QY 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKOGGTSITVRNATFTMARSDPT 660
 DB 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKOGGTSITVRNATFTMARSDPT 660
 QY 661 LINGITFSPREGALVAVVQVCGCKSSILSLALAEKDEKSHVAKGSVAVYPOQAMOND 720
 DB 661 LINGITFSPREGALVAVVQVCGCKSSILSLALAEKDEKSHVAKGSVAVYPOQAMOND 720
 QY 721 SLRENILFGCOLPEPYSVIOACALLPDLIELPSGDRTEGEGVNLSSGQKORVSLAR 780
 DB 721 SLRENILFGCOLPEPYSVIOACALLPDLIELPSGDRTEGEGVNLSSGQKORVSLAR 780
 QY 781 AVYSNADIYLFDDPLSAVDAAVGHGHIENVIQPKMLKNKTRILVTHSMSTLPQVDYIIV 840
 DB 781 AVYSNADIYLFDDPLSAVDAAVGHGHIENVIQPKMLKNKTRILVTHSMSTLPQVDYIIV 840
 QY 841 MSGGKISEMGSYQOLLARDAFAFLRTYASTBOBDAEENGVTGVSQPGKEAQMENGK 900
 DB 841 MSGGKISEMGSYQOLLARDAFAFLRTYASTBOBDAEENGVTGVSQPGKEAQMENGK 900
 QY 901 LVYDSAGKOLQROLSSSSSYSGDISRHHNSTAEIQAKEAKEETWKLMEADKAQTGOVKL 960
 DB 901 LVYDSAGKOLQROLSSSSSYSGDISRHHNSTAEIQAKEAKEETWKLMEADKAQTGOVKL 960
 QY 961 SVYVDYMKAIQGLFISFLSIFLPMCNHVSALASNTWLSIWDPIVNGTOEHTKYRLSVYG 1020
 DB 961 SVYVDYMKAIQGLFISFLSIFLPMCNHVSALASNTWLSIWDPIVNGTOEHTKYRLSVYG 1020

QY 1021 ALGISOGIAVGVSAVSGIGILLASRCIHDVLIHSILSRPMSFFERTPSGLVNRFSKEL 1080
 DB 1021 ALGISOGIAVGVSAVSGIGILLASRCIHDVLIHSILSRPMSFFERTPSGLVNRFSKEL 1080
 QY 1081 DTVDMSIPEVIMKEMGSLFNVIAGACIVILLATPIAIIIPPLGIYFVFQRFYVASSROL 1140
 DB 1081 DTVDMSIPEVIMKEMGSLFNVIAGACIVILLATPIAIIIPPLGIYFVFQRFYVASSROL 1140
 QY 1141 KRLSVSRSPYSHFNFTLGVSVIRAFEBQERTIHOSDLKVDENOKAYPSIVANRWLA 1200
 DB 1141 KRLSVSRSPYSHFNFTLGVSVIRAFEBQERTIHOSDLKVDENOKAYPSIVANRWLA 1200
 QY 1201 VRLCEVGNCTIVFPAALFAVISRHSLSAGLVGLSVSYSLQVTTYNMLVRMSSEMETNIVA 1260
 DB 1201 VRLCEVGNCTIVFPAALFAVISRHSLSAGLVGLSVSYSLQVTTYNMLVRMSSEMETNIVA 1260
 QY 1261 VERLKEYSETEKEAPMOIOTETAPSSMPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
 DB 1261 VERLKEYSETEKEAPMOIOTETAPSSMPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
 QY 1321 EKVGIVRTGAGKSLTLGLFRINESABGEIINDGINIAKIGLHDLARKTITITQDPVLF 1380
 DB 1321 EKVGIVRTGAGKSLTLGLFRINESABGEIINDGINIAKIGLHDLARKTITITQDPVLF 1380
 QY 1381 SGSLRMLDPEPSQVSDSEVWTSLELAHKOPVSALPKLHECAGGSGNLSVGOROLVCL 1440
 DB 1381 SGSLRMLDPEPSQVSDSEVWTSLELAHKOPVSALPKLHECAGGSGNLSVGOROLVCL 1440
 QY 1441 ABALLRKTILVDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDTYRIVL 1500
 DB 1441 ABALLRKTILVDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDTYRIVL 1500
 QY 1501 DKGEIOEYGA PSDILOQRGLFYSMADKAGLV 1531
 DB 1501 DKGEIOEYGA PSDILOQRGLFYSMADKAGLV 1531

RESULT 14
 US-08-463-092B-6
 Sequence 6, Application US/08463092B
 Patent No. 5766880
 GENERAL INFORMATION:
 APPLICANT: Cole, Susan P. C.
 TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
 STREET: Queen's University at Kingston
 CITY: Kingston
 STATE: Ontario
 COUNTRY: CANADA
 ZIP: K7L 3N6
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/463,092B
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/966,923
 FILING DATE: 27-OCT-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/025,340
 FILING DATE: 8-MAR-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/141,893

FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Mernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-092B-6

Query Match 71.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFCAGSDGSDPLMDNNTWNTSNPDFTKCFONTVLWVWPCFYLMACFPFFYLYLSRH 60
DB 1 MALRSFCAGSDGSDPLMDNNTWNTSNPDFTKCFONTVLWVWPCFYLMACFPFFYLYLSRH 60
QY 61 DRGYIOMPLANKTKTALGFLMTIVCMADLFYSFMRSGVLAAPVLVSPITLLGITTLA 120
DB 61 DRGYIOMPLANKTKTALGFLMTIVCMADLFYSFMRSGVLAAPVLVSPITLLGITTLA 120
QY 121 TFLDLERRKGVQSSGIMLTFMLVALVCALATLRSKIMTALKEADAOVDFRDTYYVES 180
DB 121 TFLDLERRKGVQSSGIMLTFMLVALVCALATLRSKIMTALKEADAOVDFRDTYYVES 180
QY 181 LLLIQLVLSCFSDRSLPSETIHDNPNCPBSSASFLSRTTFWMTGLIVRGYROPLEGGD 240
DB 181 LVLVLQVLSCFSDRSLPSETIHDNPNCPBSSASFLSRTTFWMTGLIVRGYROPLEGGD 240
QY 241 LMSLNKERTSEOVNVLVKNMKKECAKTRKQPKVYVYS-SKDPAPKSSSKVDANEVEYA 299
DB 241 LMSLNKERTSEOVNVLVKNMKKECAKTRKQPKVYVYS-SKDPAPKSSSKVDANEVEYA 299
QY 300 LIVSPKEMPSLRFKLYKTRGPFLMSFPFKAIHDMFSGPOLIKLIFVNDTKAP 359
DB 301 LIVSPKEMPSLRFKLYKTRGPFLMSFPFKAIHDMFSGPOLIKLIFVNDTKAP 359
QY 360 DMQGFYTVLLEFVTAQLTLVLAHQYFHLICFVSGMRIKTAVIGAVVRKALVITNSARKSST 419
DB 361 DMQGFYTVLLEFVTAQLTLVLAHQYFHLICFVSGMRIKTAVIGAVVRKALVITNSARKSST 419
QY 420 VGEIYNLMSVDAQRPMDLATYINMTWSAPLOVITALLYLMLNLGSPVLAAGVAVMYLMPV 479
DB 421 VGEIYNLMSVDAQRPMDLATYINMTWSAPLOVITALLYLMLNLGSPVLAAGVAVMYLMPV 479
QY 480 NAVVAMKTKTYOVAMHMSKDNRIKIMNEIINGIKYLKLYAMELAKFDKYLAIROBELVYL 539
DB 481 NAVVAMKTKTYOVAMHMSKDNRIKIMNEIINGIKYLKLYAMELAKFDKYLAIROBELVYL 539
QY 540 KKSAYLSAVGTFTWCTPPLVALCTPAVYVITIDENNILDAQTAFLVSLAFNLIRPPLNLT 599
DB 541 KKSAYLSAVGTFTWCTPPLVALCTPAVYVITIDENNILDAQTAFLVSLAFNLIRPPLNLT 599
QY 600 PMVTSISYVQSVSLKRLRIFLSHEBELPDSIRRPVYCGGNSITVNNATVTMARSDPP 659
DB 601 PMVTSISYVQSVSLKRLRIFLSHEBELPDSIRRPVYCGGNSITVNNATVTMARSDPP 659
QY 660 TLNGITFSIPGALVAVVGVCGKSSLLSALLAEMDKVEGHVALKGSVAVYPOQAMION 719
DB 660 TLNGITFSIPGALVAVVGVCGKSSLLSALLAEMDKVEGHVALKGSVAVYPOQAMION 719

QY 720 DSLRENTLFGCQLEPPRYRYVQAALLPDLEILIPSGDRTEIGRKGVNLSGGQQRVSLA 779
DB 720 DSLRENTLFGHPLQOENRYKALVMEACALLPDLEILIPSGDRTEIGRKGVNLSGGQQRVSLA 779
QY 780 RAVYSNADIVYFDDPLSAVDAHVGKHLFENVIGPKMLKNTKRLIYTHSMYSLPOVDYI 839
DB 780 RAVYSNADIVYFDDPLSAVDAHVGKHLFENVIGPKMLKNTKRLIYTHSMYSLPOVDYI 839
QY 840 VMSGKXISEMGSYOELLARDGAFAPLRTYASTEOBDAENGYGVSGPKCKAKOMENG 899
DB 840 VMSGKXISEMGSYOELLARDGAFAPLRTYASTEOBDAENGYGVSGPKCKAKOMENG 899
QY 900 MLVTDASGKOLQOLSSSSSYSGDISRHNSSTAELQALBAKGETWKLMEADKQOTGVK 959
DB 900 MLVTDVWKGKHLQRLHLSNSSSHSGSDTSQOHSIAELQAKA-KERTWKLMEADKQOTGVQ 955
QY 960 LSVYWDYWKALGLFISPLSIFLFCNHNVSALASNYMLSLMTDPD-PYNGOEHKRVLSY 1018
DB 960 LSVYWNWTKALGLITPLSLIFLFCNHNVSALASNYMLSLMTDPDPPVNGTOANRFLSV 1015
QY 1019 YGALGISQGLAVFGYSMAVSIIGJLASRCIHDVLDHLSILSPMSFFERTPSGNTVNRPSK 1078
DB 1016 YGALGITIQGALIFGYSMAVSIIGJFASRBLHLDLVNVLSPMSFFERTPSGNTVNRPSK 1075
QY 1079 ELDTVDYSKIPREVIVMFGSLFNVTGACVILLATPILAIIPYGLIYFFQRYVYASSR 1138
DB 1076 ELDTVDYSKIPREVIVMFGSLFNVTGACVILLATPILAIIPYGLIYFFQRYVYASSR 1135
QY 1139 QLRKLESYRSPPVYSHNETLIGSVIRAPEQERFIHQSDLKVDENOKAYPSIVANRW 1198
DB 1136 QLRKLESYRSPPVYSHNETLIGSVIRAPEQERFIHQSDLKVDENOKAYPSIVANRW 1195
QY 1199 LAVRELCVGNICVIFPALFAVISRHSLSAGLVGSVSIQITVYIMLVMSSEMETNI 1258
DB 1196 LAVRELCVGNICVIFPALFAVISRHSLSAGLVGSVSIQITVYIMLVMSSEMETNI 1255
QY 1259 VAVERLKEYSTEXBAPMOIQETAPPSWPQVGVVERPNYTLKLRREDLFTVLRHINVTIN 1318
DB 1256 VAVERLKEYSTEXBAPMOIQETAPPSWPQVGVVERPNYTLKLRREDLFTVLRHINVTIE 1315
QY 1319 GGEKVGIVGRTGACKSSITGLFRINSAGEEIIIDGNIAKIGLHDIRFKITIIIPDPV 1378
DB 1316 GGEKVGIVGRTGACKSSITGLFRINSAGEEIIIDGNIAKIGLHDIRFKITIIIPDPV 1375
QY 1379 LFGSGLRKNLDPFQYSDDEEYVTSLELAHLKDIFYSALPDKLDHCAEGENLSVGORQV 1438
DB 1376 LFGSGLRKNLDPFQYSDDEEYVTSLELAHLKDIFYSALPDKLDHCAEGENLSVGORQV 1435
QY 1439 CLARALLKTKILVDEATAVADLETDDLIOSTRTOFEDCTVLTIAHRLNTINDYRVI 1498
DB 1436 CLARALLKTKILVDEATAVADLETDDLIOSTRTOFEDCTVLTIAHRLNTINDYRVI 1495
QY 1499 VLDKGEIOEGAPSDLLQORGLFYSMADAGLV 1531
DB 1496 VLDKGEVARECAPSELLQORGIIFYSMADAGLV 1528

RESULT 15
US-08-462-109A-6
Sequence 6, Application US/08462109A
Patent No. 5882875
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,109A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deconti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-6

Query Match 71.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRFGSADGSDPLMNMNTWNTSNDFTKCFONTVLVWPCFYLMACEPFYLYSRH 60
DB 1 MALRFSFCSADGSDPLMNMNTWNTSNDFTKCFONTVLVWPCFYLMACEPFYLYSRH 60
QY 61 DRGIIQNTPLNKTATAGFLMITYCMADLFYSFWERSRGITLAVPLVSPILGITTLLA 120
DB 61 DRGIIQNTPLNKTATAGFLMITYCMADLFYSFWERSRGITLAVPLVSPILGITTLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMLVALYCALAIIKSKIMTALKEDAOYDLFRDITFYVFS 180
DB 121 TFLIQLERRKGVSSGIMLTFMLVALYCALAIIKSKIMTALKEDAOYDLFRDITFYVFS 180
QY 181 LLLIQLVLSGFSRSLPFSSTTHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSGFSRSLPFSSTTHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPUYKMKKECACTKROPYVWVS-SKDPAQPKESSVNDNEVEA 299
DB 241 LMSLNKEDTSEOVVPUYKMKKECACTKROPYVWVS-SKDPAQPKESSVNDNEVEA 299
QY 300 LIVESPOKEMNPISLFKLYKTFPGPYFLMSFFKAIHDLMPFSGOILKLIKFVNDTKAP 359
DB 300 LIVESPOKEMNPISLFKLYKTFPGPYFLMSFFKAIHDLMPFSGOILKLIKFVNDTKAP 359
QY 360 DWQGYFTVLLFTYACIQTLLVLAHOYFHI CPVSGMRITAVI GAVYRRALVITNSARKST 419
DB 360 DWQGYFTVLLFTYACIQTLLVLAHOYFHI CPVSGMRITAVI GAVYRRALVITNSARKST 419
QY 420 VGEIIVNLSYDAQSFMDLATYINMIWGAPOVITLALVLMNIGPSVLGAGAVWVLMVAV 479
DB 420 VGEIIVNLSYDAQSFMDLATYINMIWGAPOVITLALVLMNIGPSVLGAGAVWVLMVAV 479
QY 480 NAYVAMKTKTYQVAHMSKONRIKLMEIINGIKVLYKLXAMELAFKQVLAIRQEEHKLVL 539
DB 480 NAYVAMKTKTYQVAHMSKONRIKLMEIINGIKVLYKLXAMELAFKQVLAIRQEEHKLVL 539
QY 540 KKSAYLSANGFTWVCTPELVALCTFAVVTYIDENNILDAQTAFLALFNILRPEPIL 599
DB 540 KKSAYLSANGFTWVCTPELVALCTFAVVTYIDENNILDAQTAFLALFNILRPEPIL 599
QY 541 KKSAYLAAGFTWVCTPELVALCTFAVVTYIDENNILDAQTAFLALFNILRPEPIL 600
DB 541 KKSAYLAAGFTWVCTPELVALCTFAVVTYIDENNILDAQTAFLALFNILRPEPIL 600
QY 600 PMVSSIVQASVSLKRIIFLSHELEPPDSIERPVVDGGTNSITRNATFTNARSDP 659
DB 600 PMVSSIVQASVSLKRIIFLSHELEPPDSIERPVVDGGTNSITRNATFTNARSDP 659
QY 660 TUNGITFISIEGALVAVVGVCGKSSILSALAEKMDKVEGHVAKGSVAVVPOQAMION 719
DB 660 TUNGITFISIEGALVAVVGVCGKSSILSALAEKMDKVEGHVAKGSVAVVPOQAMION 719
QY 720 DSIARENILFGHPIDENYKKAEMKCALLPDLIETIPSDRTEIGKGVNLSSGQKORVSLA 779
DB 720 DSIARENILFGHPIDENYKKAEMKCALLPDLIETIPSDRTEIGKGVNLSSGQKORVSLA 779
QY 780 RAYVSNADIVLFDPLSAVDAHVGKHI FENVIGKGLKNTKRLVTHMSYLPQVDTYI 839
DB 780 RAYVSNADIVLFDPLSAVDAHVGKHI FENVIGKGLKNTKRLVTHMSYLPQVDTYI 839
QY 840 VMSGGKISEMGSYGELLARDGAPAEFLRTYASTEQEBDAEENGTVGVSGPKKAKOMENG 899
DB 840 VMSGGKISEMGSYGELLARDGAPAEFLRTYASTEQEBDAEENGTVGVSGPKKAKOMENG 899
QY 900 MLVTDGAKGOLOROLSSSSSYSGDISHHNSYALQAEAKKETWKLMEADKQOTGVK 959
DB 900 MLVTDGAKGOLOROLSSSSSYSGDISHHNSYALQAEAKKETWKLMEADKQOTGVK 959
QY 960 LSVYWDYMKAIIGFISFLSIFLFCMCHVSLASVYMLSLMTDD-PINGTOEHKVVLSV 1018
DB 960 LSVYWDYMKAIIGFISFLSIFLFCMCHVSLASVYMLSLMTDD-PINGTOEHKVVLSV 1018
QY 996 LSVYWDYMKAIIGFISFLSIFLFCMCHVSLASVYMLSLMTDD-PINGTOEHKVVLSV 1015
DB 996 LSVYWDYMKAIIGFISFLSIFLFCMCHVSLASVYMLSLMTDD-PINGTOEHKVVLSV 1015
QY 1019 YGALGISQGIIVAGSYNAVSIIGLILASRCVHDLHLSILSPMSFFERTSGNLYNFSK 1078
DB 1019 YGALGISQGIIVAGSYNAVSIIGLILASRCVHDLHLSILSPMSFFERTSGNLYNFSK 1078
QY 1079 ELDTVDSMIDPEVIMKMGSLFPNVIACIYVLLATPIAIIIPGLIYFFVQRYVASSR 1138
DB 1079 ELDTVDSMIDPEVIMKMGSLFPNVIACIYVLLATPIAIIIPGLIYFFVQRYVASSR 1138
QY 1133 QLRLESVSRSPPVYSHNETLLGVSVIRAREBERFTHQSDKVDENOKAYYPSIVANRW 1198
DB 1133 QLRLESVSRSPPVYSHNETLLGVSVIRAREBERFTHQSDKVDENOKAYYPSIVANRW 1198
QY 1199 LAVALRECVGNCIYLFALFAVISRHSIAGLVGSIYSIQVITVYLMVLRMSSEMETNI 1258
DB 1199 LAVALRECVGNCIYLFALFAVISRHSIAGLVGSIYSIQVITVYLMVLRMSSEMETNI 1258
QY 1259 VAVERLKEYSETEKEAPWQIOETAPPSGWPQVGRVEFRNVCYLRREDLDFVLRHINTIN 1318
DB 1259 VAVERLKEYSETEKEAPWQIOETAPPSGWPQVGRVEFRNVCYLRREDLDFVLRHINTIN 1318
QY 1319 GGEKVGIVGRGTGAKSSTLTGLFRINESAGCEIITDGVNIAKIGLHNRKRTIIPDDV 1378
DB 1319 GGEKVGIVGRGTGAKSSTLTGLFRINESAGCEIITDGVNIAKIGLHNRKRTIIPDDV 1378
QY 1379 LPSGSLAMNIDPFSQYSDSEEVWTSLELAHLKDFVSALPDLKDBHCAEGENLSVGQROLY 1438
DB 1379 LPSGSLAMNIDPFSQYSDSEEVWTSLELAHLKDFVSALPDLKDBHCAEGENLSVGQROLY 1438
QY 1436 CLARALIRKTKIIVLEDEATAVDETNLQSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1495
DB 1436 CLARALIRKTKIIVLEDEATAVDETNLQSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1495
QY 1499 VLDKGRIOEYGAAPSDDLQGRGLFYMAKDAGLV 1531
DB 1499 VLDKGRIOEYGAAPSDDLQGRGLFYMAKDAGLV 1531
QY 1528 VLDKGRVBCGAPSELDQGRGIFYMAKDAGLV 1528
DB 1528 VLDKGRVBCGAPSELDQGRGIFYMAKDAGLV 1528

Search completed: December 15, 2005, 15:49:42
Job time : 48.351 secs

This page is blank.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:00 ; Search time 156.846 Seconds
(without alignments)
5037.528 Million cell updates/sec

Title: US-10-665-283-4

Perfect score: 9734

Sequence: 1 MALRGFCSADGSDPLMDMNV.....NTIKVPTPLCTARQLDEDRS 1891

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9734	100.0	1891	US-10-665-283-4	Sequence 4, Appl1
2	9734	100.0	1927	US-10-665-283-1	Sequence 1, Appl1
3	9734	100.0	1927	US-10-665-283-6	Sequence 6, Appl1
4	9704	99.7	1947	US-10-665-283-8	Sequence 8, Appl1
5	7860	80.7	1531	US-09-939-853A-86	Sequence 86, Appl1
6	7860	80.7	1531	US-10-667-891-6	Sequence 6, Appl1
7	7860	80.7	1531	US-10-807-466-6	Sequence 6, Appl1
8	7860	80.7	1531	US-10-484-577-678	Sequence 678, Appl1
9	7860	80.7	1531	US-10-889-503-19	Sequence 19, Appl1
10	7860	80.7	1531	US-10-756-149-5033	Sequence 5033, Appl1
11	7769	79.8	1515	US-09-939-853A-87	Sequence 87, Appl1
12	7763	79.8	1515	US-10-618-281-42	Sequence 42, Appl1
13	7037	72.3	1388	US-10-408-765A-1718	Sequence 1718, Appl1
14	7002.5	71.9	1528	US-09-939-853A-88	Sequence 88, Appl1
15	6068.5	62.3	1303	US-10-618-281-63	Sequence 63, Appl1
16	4697.5	48.3	1875	US-10-665-283-5	Sequence 5, Appl1
17	4697.5	48.3	1911	US-10-665-283-2	Sequence 2, Appl1
18	4667.5	48.0	1895	US-10-665-283-9	Sequence 9, Appl1
19	4667.5	48.0	1901	US-10-665-283-10	Sequence 10, Appl1
20	4491.5	46.1	1527	US-09-939-853A-83	Sequence 83, Appl1
21	4491.5	46.1	1527	US-09-939-853A-84	Sequence 84, Appl1
22	4491.5	46.1	1527	US-10-235-027-1342	Sequence 1342, Appl1
23	4487.5	46.1	1527	US-10-889-503-6	Sequence 6, Appl1
24	4487.5	46.1	1530	US-10-889-503-33	Sequence 33, Appl1
25	4479.5	45.2	1530	US-10-618-281-55	Sequence 55, Appl1
26	4396	45.2	1522	US-09-939-853A-85	Sequence 85, Appl1
27	3966	40.7	1905	US-10-665-283-3	Sequence 3, Appl1

28	3966	40.7	1941	US-10-665-283-12	Sequence 12, Appl1
29	3936	40.4	1961	US-10-665-283-7	Sequence 7, Appl1
30	3879.5	39.9	1548	US-10-807-466-28	Sequence 28, Appl1
31	3860.5	39.7	1548	US-10-667-891-1	Sequence 1, Appl1
32	3860.5	39.7	1548	US-10-807-466-1	Sequence 1, Appl1
33	3846	39.5	1538	US-10-807-466-23	Sequence 23, Appl1
34	3841	39.5	1538	US-10-807-466-25	Sequence 25, Appl1
35	3835	39.4	1538	US-10-807-466-24	Sequence 24, Appl1
36	3817.5	39.2	1539	US-10-807-466-26	Sequence 26, Appl1
37	3817	39.2	1591	US-11-097-143-13266	Sequence 13266, A
38	3797	39.0	1538	US-10-807-466-27	Sequence 27, Appl1
39	3779	38.8	1538	US-10-807-466-16	Sequence 16, Appl1
40	3774	38.8	1538	US-10-807-466-18	Sequence 18, Appl1
41	3773	38.8	1538	US-10-807-466-22	Sequence 22, Appl1
42	3768	38.7	1538	US-10-807-466-17	Sequence 17, Appl1
43	3754	38.6	1538	US-10-807-466-19	Sequence 19, Appl1
44	3750.5	38.5	1539	US-10-807-466-20	Sequence 20, Appl1
45	3733	38.4	1547	US-11-097-143-25248	Sequence 25248, A

ALIGNMENTS

RESULT 1		US-10-665-283-4
Sequence 4, Application US/10665283		
Publication No. US20050063989A1		
GENERAL INFORMATION:		
APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE		
APPLICANT: DERAND, Renaud		
APPLICANT: GARCIA, Elisabeth		
APPLICANT: PROST, Anne-Lise		
APPLICANT: REVILLAUD, Jean		
APPLICANT: VIVAUDOU, Michel		
TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF		
FILE REFERENCE: P263US104		
CURRENT APPLICATION NUMBER: US/10/665,283		
CURRENT FILING DATE: 2003-09-22		
NUMBER OF SEQ ID NOS: 25		
SOFTWARE: PatentIn version 3.1		
SEQ ID NO 4		
LENGTH: 1891		
TYPE: PRT		
ORGANISM: Homo sapiens		
US-10-665-283-4		
Query Match		100.0%; Score 9734; DB 5; Length 1891;
Best Local Similarity		100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MALRGFCSADGSDPLMDMNVNTNTSNPDFTKCFONTLVVWPCEYIMACFPFFLYLSRH 60	
DB	1 MALRGFCSADGSDPLMDMNVNTNTSNPDFTKCFONTLVVWPCEYIMACFPFFLYLSRH 60	
QY	61 DRGIQMTPLNKTKTALGELMTVCADLTFYSEWERSRGIFLAPVPLVSTLIGITTLA 120	
DB	61 DRGIQMTPLNKTKTALGELMTVCADLTFYSEWERSRGIFLAPVPLVSTLIGITTLA 120	
QY	121 TPLIQERRRGVSSGIMLTFMLVALYCALAIRSKIMTALKXDAQVDLFRDITFYVFS 180	
DB	121 TPLIQERRRGVSSGIMLTFMLVALYCALAIRSKIMTALKXDAQVDLFRDITFYVFS 180	
QY	181 LLLIQVLSCFSRSPPLFSETIHDPNPCPESSASFISRIFFWMTGLIVGYQPLEGSD 240	
DB	181 LLLIQVLSCFSRSPPLFSETIHDPNPCPESSASFISRIFFWMTGLIVGYQPLEGSD 240	
QY	241 LMSLNKEDTSEQVVPVVKMKKECAKTRQPKVYVSSQDPAPQPESSAVDANEVEAL 300	
DB	241 LMSLNKEDTSEQVVPVVKMKKECAKTRQPKVYVSSQDPAPQPESSAVDANEVEAL 300	
QY	301 IVKSPQKKNPSPLEKVLTKTGGPYFLMSFFPKAIHDLAMFSGPQILKLTKFVNDTKAPD 360	
DB	301 IVKSPQKKNPSPLEKVLTKTGGPYFLMSFFPKAIHDLAMFSGPQILKLTKFVNDTKAPD 360	

QY 361 MGGYFTYVLLFVTAQLOTLVLHOFPHICFVSGBMRTKAVIGAVYRKALVITNSARKSSTV 420
DB 361 MGGYFTYVLLFVTAQLOTLVLHOFPHICFVSGBMRTKAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAQREMDLATYINMITSAPLOVITLALYLMNLGSGVLAVAVMVLMPVN 480
DB 421 GEIYNLMSVDAQREMDLATYINMITSAPLOVITLALYLMNLGSGVLAVAVMVLMPVN 480
QY 481 AVAMAKTKTYOVAMHKSNDNRILKLMNEITLNGIKVLKLYAMELAFDKYLAIROEELKYLK 540
DB 481 AVAMAKTKTYOVAMHKSNDNRILKLMNEITLNGIKVLKLYAMELAFDKYLAIROEELKYLK 540
QY 541 KSAVLSAVGTFTWCTPELVALCTFAVYVTTIDENNITLDAQTAFYSLAFENILREPPLNLP 600
DB 541 KSAVLSAVGTFTWCTPELVALCTFAVYVTTIDENNITLDAQTAFYSLAFENILREPPLNLP 600
QY 601 MVISSIVQASVSLKRLRILFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
DB 601 MVISSIVQASVSLKRLRILFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
QY 661 LINGITFSIPEGALVAVVQVCGKSSLLSALLAEMDKYGHVAIKGSVAYVPOQAMIOND 720
DB 661 LINGITFSIPEGALVAVVQVCGKSSLLSALLAEMDKYGHVAIKGSVAYVPOQAMIOND 720
QY 721 SLRENILFGCQLEBPYRSVIOACALLPDLILPSGDRTIEIGKGVNLSGGQKQVSLAR 780
DB 721 SLRENILFGCQLEBPYRSVIOACALLPDLILPSGDRTIEIGKGVNLSGGQKQVSLAR 780
QY 781 AVYENADITLFPDDPLSANDAVGKHIFENVIGPKMKLNKTRILVTHSMSTLPQVDVITV 840
DB 781 AVYENADITLFPDDPLSANDAVGKHIFENVIGPKMKLNKTRILVTHSMSTLPQVDVITV 840
QY 841 MSGGKISEMGSYOELLARDGAPAEFLRTYASTEOBDAEENGCVGSGPGKBAKMEKMG 900
DB 841 MSGGKISEMGSYOELLARDGAPAEFLRTYASTEOBDAEENGCVGSGPGKBAKMEKMG 900
QY 901 LVTDSAGKOLOROLSSSSSSSGDISSRHNSTAELOKAKKEETKLMKLEADKAQOTGYKL 960
DB 901 LVTDSAGKOLOROLSSSSSSSGDISSRHNSTAELOKAKKEETKLMKLEADKAQOTGYKL 960
QY 961 SVYNDYKAKILGFLISFLSIFLPMCHNVASALSNVLSMTDPIYNGOETHKVLSTYGG 1020
DB 961 SVYNDYKAKILGFLISFLSIFLPMCHNVASALSNVLSMTDPIYNGOETHKVLSTYGG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGLILSRCHVLDLHLSILSPSFERTSGMLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGLILSRCHVLDLHLSILSPSFERTSGMLVNRFSKEL 1080
QY 1081 DTVDSMTPEVIAKMFMSGLFENVIGACIVILLATPPIAIIIPPLGLIYFFVQRFYVASSRQL 1140
DB 1081 DTVDSMTPEVIAKMFMSGLFENVIGACIVILLATPPIAIIIPPLGLIYFFVQRFYVASSRQL 1140
QY 1141 KRLFSVSRSPYVSHFNETLLGVSVIRAFEBQERFHQSDLKVDENOKAYPYSIVANRWLA 1200
DB 1141 KRLFSVSRSPYVSHFNETLLGVSVIRAFEBQERFHQSDLKVDENOKAYPYSIVANRWLA 1200
QY 1201 VRLFCVNGCIVLFPALFAVISRHSLSAGLVGSLSVSYLQVTTYTNMLVRMSEMETNTIVA 1260
DB 1201 VRLFCVNGCIVLFPALFAVISRHSLSAGLVGSLSVSYLQVTTYTNMLVRMSEMETNTIVA 1260
QY 1261 VERLKEVSETEKEAPWJOETAPPSMPQVGVFEFRNTCLARYEDLDVLRHINVTINGG 1320
DB 1261 VERLKEVSETEKEAPWJOETAPPSMPQVGVFEFRNTCLARYEDLDVLRHINVTINGG 1320
QY 1321 EKVGIIVGTGAGKSSLLGLFRINSSAGEIITIDGINIAKIGLHDLRFKRTIITIPDPVLF 1380
DB 1321 EKVGIIVGTGAGKSSLLGLFRINSSAGEIITIDGINIAKIGLHDLRFKRTIITIPDPVLF 1380
QY 1381 SGSLRMLNDPFSQYSDEEVTWSTLELAHLKDPVSLPDKLDHCCAGGENTLSVGQOLVCL 1440
DB 1381 SGSLRMLNDPFSQYSDEEVTWSTLELAHLKDPVSLPDKLDHCCAGGENTLSVGQOLVCL 1440

QY 1441 ARALLRRTKXILVDEATAVADLEFDDLIQSTIRFOEEDCVYLTIAHRLNTIMDYTRYVL 1500
DB 1441 ARALLRRTKXILVDEATAVADLEFDDLIQSTIRFOEEDCVYLTIAHRLNTIMDYTRYVL 1500
QY 1501 DKGEIOEYGA PSDLLOQRGLFYSMAKADGLVGGGGGMLSRKGIIPREYVLTRLAEPDAE 1560
DB 1501 DKGEIOEYGA PSDLLOQRGLFYSMAKADGLVGGGGGMLSRKGIIPREYVLTRLAEPDAE 1560
QY 1561 PRYTRERRARFVSKKNCNVAHKNIREQGRFLQDVFTTVLVDLKMPTHLLIFMSFLCSW 1620
DB 1561 PRYTRERRARFVSKKNCNVAHKNIREQGRFLQDVFTTVLVDLKMPTHLLIFMSFLCSW 1620
QY 1621 LLFAMVWMLIAFAHGDIAPEGIVNVCVTSIHSFSSAFLESIEVQVITIGGCGRWTEBCP 1680
DB 1621 LLFAMVWMLIAFAHGDIAPEGIVNVCVTSIHSFSSAFLESIEVQVITIGGCGRWTEBCP 1680
QY 1681 LALILILVONTVGMINAIVMGCIFMKTQAHRRPAEFLILSKHAVITLRRGLCFMRLVYG 1740
DB 1681 LALILILVONTVGMINAIVMGCIFMKTQAHRRPAEFLILSKHAVITLRRGLCFMRLVYG 1740
QY 1741 DLKSKMIIISATIHQVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAPLIIYHVIDS 1800
DB 1741 DLKSKMIIISATIHQVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAPLIIYHVIDS 1800
QY 1801 NSPLYDLAPSDLHHODLEIIVIEGVETTGITTOARTSYLADELLMGQRFVPIVAED 1860
DB 1801 NSPLYDLAPSDLHHODLEIIVIEGVETTGITTOARTSYLADELLMGQRFVPIVAED 1860
QY 1861 GRYSVDYSKRGNTIKVPTPLCTARQLEDRS 1891
DB 1861 GRYSVDYSKRGNTIKVPTPLCTARQLEDRS 1891

RESULT 2

US-10-665-283-1
; Sequence 1, Application US/10665283
; Publication No. US2005063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L' ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLAUD, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: P263US104
; CURRENT APPLICATION NUMBER: US/10/665, 283
; CURRENT FILING DATE: 2003-09-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1927
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-1
Query Match 100.0%; Score 9734; DB 5; Length 1927;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALNGFCGADSDPLMDMNVNTMNTSNDPFTKCFQNTVVLVWPCFYLMACPFYFYLSRH 60
DB 1 MALNGFCGADSDPLMDMNVNTMNTSNDPFTKCFQNTVVLVWPCFYLMACPFYFYLSRH 60
QY 61 DRGIYQMTPLANKTGTALGFLIMIVCMADLFYSFEMERSGCIPLAVFLVSPFLIGITTLA 120
DB 61 DRGIYQMTPLANKTGTALGFLIMIVCMADLFYSFEMERSGCIPLAVFLVSPFLIGITTLA 120
QY 121 TFLIQLEBRKGVSSGIMLTFWLVALVICALIISKIMTAKEDPAQVDFPDIFPYVFS 180
DB 121 TFLIQLEBRKGVSSGIMLTFWLVALVICALIISKIMTAKEDPAQVDFPDIFPYVFS 180
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPBSSASFLSRTFPMWITGLIVRGYROPLEGSD 240

Db 181 LLLIOLVSCFSDSPLFSEFTHDBNPCRESSASPLSRITFWMTGLIVRGYQGLESD 240
Qy 241 LMSLNKETSBOVAVPVLYKMKKECAKTRKOPVKVYSSKDPAPKESKVDANEVAL 300
Db 241 LMSLNKETSBOVAVPVLYKMKKECAKTRKOPVKVYSSKDPAPKESKVDANEVAL 300
Qy 301 IVKSPKEMNPSLFVLYKTFGPYFLMSFFPKAHDLMFSGPOLIKLLIFVNDTKAPD 360
Db 301 IVKSPKEMNPSLFVLYKTFGPYFLMSFFPKAHDLMFSGPOLIKLLIFVNDTKAPD 360
Qy 361 WQGFYTYLLFVTACLOTVLHOFYFHICFVSGMRKTAIVIGAVYRKALVITNSARKSTV 420
Db 361 WQGFYTYLLFVTACLOTVLHOFYFHICFVSGMRKTAIVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLSVDAQRFPDLDATYINMTWSAPLOVILALYLLMLNGSVLGAVMVLMVFN 480
Db 421 GEIVNLSVDAQRFPDLDATYINMTWSAPLOVILALYLLMLNGSVLGAVMVLMVFN 480
Qy 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLYKLYAMELAFKDKYLAIOBELKYLK 540
Db 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLYKLYAMELAFKDKYLAIOBELKYLK 540
Qy 541 KSAVLSAVGTFWCTPPLVALCTFAVYVITDENNIIDAQTAFYSLAFNILRFPNLTP 600
Db 541 KSAVLSAVGTFWCTPPLVALCTFAVYVITDENNIIDAQTAFYSLAFNILRFPNLTP 600
Qy 601 MVISSIVOASVSLKRLRIFLSHEBIEPDSIERRPYKOGGJNSITVRNATTTMARSDPT 660
Db 601 MVISSIVOASVSLKRLRIFLSHEBIEPDSIERRPYKOGGJNSITVRNATTTMARSDPT 660
Qy 661 LINGTFSPGEGALVAVVGVGCGKSSLSALLAEMDKYGHVALIKGSVAVYPOQAMOND 720
Db 661 LINGTFSPGEGALVAVVGVGCGKSSLSALLAEMDKYGHVALIKGSVAVYPOQAMOND 720
Qy 721 SLRENIIFGCOLLEBPYRSVIOACALLPDLLEILPBGDRTEIGEGVNLSSGOKRVSLAR 780
Db 721 SLRENIIFGCOLLEBPYRSVIOACALLPDLLEILPBGDRTEIGEGVNLSSGOKRVSLAR 780
Qy 781 AVYSNADIYLFDDPLSAVDHAVGHIFENVIGPKMKLNKTRILVTHSMYSLPOVDVITV 840
Db 781 AVYSNADIYLFDDPLSAVDHAVGHIFENVIGPKMKLNKTRILVTHSMYSLPOVDVITV 840
Qy 841 MSGKISGEMGYOELLARDGAFAPLRTYASTBOGDAENGYGVSGPGKAAQOMENGM 900
Db 841 MSGKISGEMGYOELLARDGAFAPLRTYASTBOGDAENGYGVSGPGKAAQOMENGM 900
Qy 901 LVTDSAGKOLOROSSSSSYSGDISRHNSYAELOKAEAKKEETWKLMEADKAOTGOVKL 960
Db 901 LVTDSAGKOLOROSSSSSYSGDISRHNSYAELOKAEAKKEETWKLMEADKAOTGOVKL 960
Qy 961 SVYWDYKAIGLFISFLSIFLFCMNVHVSALASNYLSMTDPIVNGTOEHTKRLSVYG 1020
Db 961 SVYWDYKAIGLFISFLSIFLFCMNVHVSALASNYLSMTDPIVNGTOEHTKRLSVYG 1020
Qy 1021 ALGISOGIAGVGYMAVSIIGIILASRCHVULHLSIILSPMSFFERTPGCVLVNRFSEKL 1080
Db 1021 ALGISOGIAGVGYMAVSIIGIILASRCHVULHLSIILSPMSFFERTPGCVLVNRFSEKL 1080
Qy 1081 DTVDSMPEVIKMGSLFNVIGACIYILLATPIAIIIPPLGIIYFPVORFYVASSROL 1140
Db 1081 DTVDSMPEVIKMGSLFNVIGACIYILLATPIAIIIPPLGIIYFPVORFYVASSROL 1140
Qy 1141 KRLESVRSRPSVSHFNELLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
Db 1141 KRLESVRSRPSVSHFNELLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
Qy 1201 VRLCEVNCIYLFALFAVVISRHSLSAGIVLSYSIQVTTYANMLVYKSSMETNIVA 1260
Db 1201 VRLCEVNCIYLFALFAVVISRHSLSAGIVLSYSIQVTTYANMLVYKSSMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVEFRNYCLAREDDLPVLRIHINVTING 1320
Db 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVEFRNYCLAREDDLPVLRIHINVTING 1320

Db 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVEFRNYCLAREDDLPVLRIHINVTING 1320
Qy 1321 EKVGIVERTAGKXSLTGLFRINESAGEIINDGINIAKIGHDLRFTITTIPODPLF 1380
Db 1321 EKVGIVERTAGKXSLTGLFRINESAGEIINDGINIAKIGHDLRFTITTIPODPLF 1380
Qy 1381 SGSLRMLDPFSQYSDDEWTSLELAHKDFVSALPKLDHECAGEENLSVGOROLVCL 1440
Db 1381 SGSLRMLDPFSQYSDDEWTSLELAHKDFVSALPKLDHECAGEENLSVGOROLVCL 1440
Qy 1441 ARALLRKTILVDEATAVADLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
Db 1441 ARALLRKTILVDEATAVADLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
Qy 1501 DKGEIOEYGAPSDILQORGLFYSAKAKAGLVGGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
Db 1501 DKGEIOEYGAPSDILQORGLFYSAKAKAGLVGGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
Qy 1561 PRYTRRRRARFVSKKNCNVAHKNIREOGRFLQDVFTLLVDLKMPTLLIFTMSPLCSW 1620
Db 1561 PRYTRRRRARFVSKKNCNVAHKNIREOGRFLQDVFTLLVDLKMPTLLIFTMSPLCSW 1620
Qy 1621 LLFAMVWMLIAFAHGLDAPGSGTVPCVTSIHSFSAFLSIEVQVITIGFGRMVTBEP 1680
Db 1621 LLFAMVWMLIAFAHGLDAPGSGTVPCVTSIHSFSAFLSIEVQVITIGFGRMVTBEP 1680
Qy 1681 LAILILVONIIVGMINAIMLCTFMKTAQHRAEFTLISKAIVITLRHGRCLCPMLRVG 1740
Db 1681 LAILILVONIIVGMINAIMLCTFMKTAQHRAEFTLISKAIVITLRHGRCLCPMLRVG 1740
Qy 1741 DLKSMITISATIHQVVRKTTSPGEVVPVPHOVDIPENGVGNGIFLVAPLIYHYIDS 1800
Db 1741 DLKSMITISATIHQVVRKTTSPGEVVPVPHOVDIPENGVGNGIFLVAPLIYHYIDS 1800
Qy 1801 NSPLVLDAPSDLHHODLEIIVILEGVETTGITTOARTSYLADEIIMGORFVPIVAED 1860
Db 1801 NSPLVLDAPSDLHHODLEIIVILEGVETTGITTOARTSYLADEIIMGORFVPIVAED 1860
Qy 1861 GRYSVDYSKFGNTIKVPTPLCTARQLEDEDS 1891
Db 1861 GRYSVDYSKFGNTIKVPTPLCTARQLEDEDS 1891

RESULT 3
US-10-665-283-6
; Sequence 6, Application US/10665283
; Publication No. US20050063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLAUD, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OR INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREO
; FILE REFERENCE: P26305104
; CURRENT APPLICATION NUMBER: US/10/665,283
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1927
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-6

Query Match 100.0%; Score 9734; DB 5; Length 1927;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDMNTVNTSNDPFTKCFONTLVVWPCFYLMACPPFFLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTVNTSNDPFTKCFONTLVVWPCFYLMACPPFFLYLSRH 60

QY 61 DRGYIOMTPLNKRTALGFLIMIVCMADLFYSFEMERSRGIPLAPVFLVSPILLGITTLA 120
Db 61 DRGYIOMTPLNKRTALGFLIMIVCMADLFYSFEMERSRGIPLAPVFLVSPILLGITTLA 120
QY 121 TFLIOLERRKVOSSGIMLTWLVALVCAALILSKINTALKEADAVDLPFDITFYVFS 180
Db 121 TFLIOLERRKVOSSGIMLTWLVALVCAALILSKINTALKEADAVDLPFDITFYVFS 180
QY 181 LLLIOLVASCSDSPLESETIHDNPPCRESSASPLRITFMWITGLIVRGROPLBSSD 240
Db 181 LLLIOLVASCSDSPLESETIHDNPPCRESSASPLRITFMWITGLIVRGROPLBSSD 240
QY 241 LMSLINKEDTSBOVVPVIVKMKKECAKTRKQPVKVVVSSKDPAPCKESSKYDANEVVAL 300
Db 241 LMSLINKEDTSBOVVPVIVKMKKECAKTRKQPVKVVVSSKDPAPCKESSKYDANEVVAL 300
QY 301 IVKSPCKEMNPSLFRVLYKTEGPFYFMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPD 360
Db 301 IVKSPCKEMNPSLFRVLYKTEGPFYFMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPD 360
QY 361 WQGFYTYVLFTVACIOFLVHOYFHICFVSGMRKTPVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYTYVLFTVACIOFLVHOYFHICFVSGMRKTPVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAQRFMDLATYINMIMSAPLOVILALYLLMLNLSVVLGAVAVMLMVPVN 480
Db 421 GEIYNLMSVDAQRFMDLATYINMIMSAPLOVILALYLLMLNLSVVLGAVAVMLMVPVN 480
QY 481 AVMAKTKTYOVAMHKSNDNRKILMNEILNGIKVLKLYAMELAFDOKYLAIRROBELKYLK 540
Db 481 AVMAKTKTYOVAMHKSNDNRKILMNEILNGIKVLKLYAMELAFDOKYLAIRROBELKYLK 540
QY 541 KSAVLASAVGTWCTPFLVALCTPAVYVTDENNILDAQTRAFVSLAFNLIRPLNLTP 600
Db 541 KSAVLASAVGTWCTPFLVALCTPAVYVTDENNILDAQTRAFVSLAFNLIRPLNLTP 600
QY 601 MVISSIVQASVSLKRLIFLSHEBELPDSIERRPYKDGCTNSITVRNATFTWASDPPT 660
Db 601 MVISSIVQASVSLKRLIFLSHEBELPDSIERRPYKDGCTNSITVRNATFTWASDPPT 660
QY 661 LNGTHFSPREGALVAVVQVCGKSSLLSALLAENDKVBGHVALIGSAVAVYPOQAMIOND 720
Db 661 LNGTHFSPREGALVAVVQVCGKSSLLSALLAENDKVBGHVALIGSAVAVYPOQAMIOND 720
QY 721 SLRENILLEGCOLEBYRSVIOACALLPDLITLPSGDRTEIGEKVNLSSGQKORVSLAR 780
Db 721 SLRENILLEGCOLEBYRSVIOACALLPDLITLPSGDRTEIGEKVNLSSGQKORVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDAAHVKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840
Db 781 AVYSNADIYLFDDPLSAVDAAHVKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840
QY 841 MSGKISMSGSYOELLARDGAPAEFLRTYASTEOBDAEENGVTGVSQPKAKEMNGM 900
Db 841 MSGKISMSGSYOELLARDGAPAEFLRTYASTEOBDAEENGVTGVSQPKAKEMNGM 900
QY 901 LVTHSAGKOLROULSSSSSYSGDISRHNSSTAELOKAKKKEBTMKLMEADAOQOYKL 960
Db 901 LVTHSAGKOLROULSSSSSYSGDISRHNSSTAELOKAKKKEBTMKLMEADAOQOYKL 960
QY 961 SVYMDYMAKIGLFIFFLSIFLPMCHVASALASNYMLMTDDPIYNGTOEHKTVLSTYVG 1020
Db 961 SVYMDYMAKIGLFIFFLSIFLPMCHVASALASNYMLMTDDPIYNGTOEHKTVLSTYVG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGILLASRCLHVDLHLSILSPSPFFERTSPGNLVNFSKEL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGILLASRCLHVDLHLSILSPSPFFERTSPGNLVNFSKEL 1080
QY 1081 DTUNSMIPEVIMKMGSLPNVIGACIVILLATPIAIIIPRGLIYFPQRYVASSRQL 1140
Db 1081 DTUNSMIPEVIMKMGSLPNVIGACIVILLATPIAIIIPRGLIYFPQRYVASSRQL 1140

QY 1141 KRLSVSRSPVYSHFNFTLLIGSVITRAPEBOERFIHOSDLKVDENQAYYPSIYANRWLA 1200
Db 1141 KRLSVSRSPVYSHFNFTLLIGSVITRAPEBOERFIHOSDLKVDENQAYYPSIYANRWLA 1200
QY 1201 VRLCEVNCCTVLPALFAVISRHSLSAGLVLSYSYLOVTTYANMLVRMSSEMETNIVA 1260
Db 1201 VRLCEVNCCTVLPALFAVISRHSLSAGLVLSYSYLOVTTYANMLVRMSSEMETNIVA 1260
QY 1261 VERLEKSETEKEAPMOIOETAPSSWPQVGRVFRNYCLRYRBDLPVLRLHINVTINGG 1320
Db 1261 VERLEKSETEKEAPMOIOETAPSSWPQVGRVFRNYCLRYRBDLPVLRLHINVTINGG 1320
QY 1321 EKVGIVRTGAGKSSLTGLPRINESABEGIIIDGINIAKIGLHDLRPKTTIIPODPYLF 1380
Db 1321 EKVGIVRTGAGKSSLTGLPRINESABEGIIIDGINIAKIGLHDLRPKTTIIPODPYLF 1380
QY 1381 SGSLRMLNDPSSQYSDSEWMTSLFLAKOPVSLPKDLHECAEGENLSVGOROLVCL 1440
Db 1381 SGSLRMLNDPSSQYSDSEWMTSLFLAKOPVSLPKDLHECAEGENLSVGOROLVCL 1440
QY 1441 ARALLRKTILVDEATRAVDLETDLLIOGTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
Db 1441 ARALLRKTILVDEATRAVDLETDLLIOGTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIOEYGAAPSDLLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPERYVLTRLAEDPAE 1560
Db 1501 DKGEIOEYGAAPSDLLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPERYVLTRLAEDPAE 1560
QY 1561 PRYTRERRARFVSKKNCNVANHNIREQRFLODVEVTTLVDLKMPHTLITFMSFLCSW 1620
Db 1561 PRYTRERRARFVSKKNCNVANHNIREQRFLODVEVTTLVDLKMPHTLITFMSFLCSW 1620
QY 1621 LLFAMVWMLIFAFAGDLPAGEGTNVCVTSIHSFSAFLFSIEVQVITIGFGRMWTEBCP 1680
Db 1621 LLFAMVWMLIFAFAGDLPAGEGTNVCVTSIHSFSAFLFSIEVQVITIGFGRMWTEBCP 1680
QY 1681 LAILILIVONIVGIMINAMLGCIEMKTAQHRAPELIFSKHAVITLRHRLCPMLRVG 1740
Db 1681 LAILILIVONIVGIMINAMLGCIEMKTAQHRAPELIFSKHAVITLRHRLCPMLRVG 1740
QY 1741 DLKRSMTISATIHNOVARKTSPGEVVPPLHOUNIIPMENGNGGCIPLVABLIIYHVIDS 1800
Db 1741 DLKRSMTISATIHNOVARKTSPGEVVPPLHOUNIIPMENGNGGCIPLVABLIIYHVIDS 1800
QY 1801 NSPLYDLAPSDLHHODLEIIVILEGVETGITTQARTSYLADEIILMGQRFVIAEED 1860
Db 1801 NSPLYDLAPSDLHHODLEIIVILEGVETGITTQARTSYLADEIILMGQRFVIAEED 1860
QY 1861 GRYSVDYSKFGNTIKVPTPLCTARQLEDERS 1891
Db 1861 GRYSVDYSKFGNTIKVPTPLCTARQLEDERS 1891

RESULT 4
US-10-665-283-8
; Sequence 8, Application US/10665283
; Publication No. US2005063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L' ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLIQU, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: F263US104
; CURRENT APPLICATION NUMBER: US/10/665, 283
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1947
; TYPE: PRT

ORGANISM: Homo sapiens
US-10-665-283-8

Query Match 99.7%; Score 9704; DB 5; Length 19471.
Beet local Similarity 99.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

1 MALRGFCAGSDSDPLMDNMVNTWNTSNPDFTKCFONTLVVWVPCFYLMACFPPYFLYLSRH 60
1 MALRGFCAGSDSDPLMDNMVNTWNTSNPDFTKCFONTLVVWVPCFYLMACFPPYFLYLSRH 60
61 DRGVIQMTPLNKTALGFLMIYVCMADLFYSFWERSRGIFLAPFVLSPTLLGTTLLA 120
61 DRGVIQMTPLNKTALGFLMIYVCMADLFYSFWERSRGIFLAPFVLSPTLLGTTLLA 120
121 TFLIQLERKGVQSSGIMLTWLVVALVCAALILSKIMTALKEDAQVDLFRDITFYVFS 180
121 TFLIQLERKGVQSSGIMLTWLVVALVCAALILSKIMTALKEDAQVDLFRDITFYVFS 180
181 LLLIQLVASCPSDRSPLSETIHDNPNCPSSASLSRITFMWITGLIVRGYRQPLEGSD 240
181 LLLIQLVASCPSDRSPLSETIHDNPNCPSSASLSRITFMWITGLIVRGYRQPLEGSD 240
241 LMSLNKEDTSEQVVPVLVKNMKKCECAKTRKQPVKVVYSSKDPAOFKSSKYDANEVVAL 300
241 LMSLNKEDTSEQVVPVLVKNMKKCECAKTRKQPVKVVYSSKDPAOFKSSKYDANEVVAL 300
301 IVKSPQKEMNPDLFKVLYKTTGPPFLMSPPFKAIHDLMPFSGPOLIKLIFVNDTKAD 360
301 IVKSPQKEMNPDLFKVLYKTTGPPFLMSPPFKAIHDLMPFSGPOLIKLIFVNDTKAD 360
361 WQGFYVTLFVTACLOTLVHOVPHLICFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
361 WQGFYVTLFVTACLOTLVHOVPHLICFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
421 GEIVNLMSVDAQRFMDLATYINMISAPLOVILALYILMLNIGPSVLGAVVMTLVAVN 480
421 GEIVNLMSVDAQRFMDLATYINMISAPLOVILALYILMLNIGPSVLGAVVMTLVAVN 480
481 AVMAKTKTYOVANHKSKDNRIKLMNELINGIKVILKYAMELAFKDYALIROBELKYLK 540
481 AVMAKTKTYOVANHKSKDNRIKLMNELINGIKVILKYAMELAFKDYALIROBELKYLK 540
541 KSAVLASVGTFTWCTPFLVALCTPAVYVTTIDENNIIIDAQTAFLSLAFNLRPPLNLP 600
541 KSAVLASVGTFTWCTPFLVALCTPAVYVTTIDENNIIIDAQTAFLSLAFNLRPPLNLP 600
601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTMARSDPT 660
601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTMARSDPT 660
661 LMGITFSPREGALVAVVGVCGKSSLLSALLAEMDKVGHVAKGSVAVYVPOQAMIOND 720
661 LMGITFSPREGALVAVVGVCGKSSLLSALLAEMDKVGHVAKGSVAVYVPOQAMIOND 720
721 SLRENIILFGCOLLEPPYRSVYQACALLPDLLEILPSGDTETIGEGVNLSSGQKORVSLAR 780
721 SLRENIILFGCOLLEPPYRSVYQACALLPDLLEILPSGDTETIGEGVNLSSGQKORVSLAR 780
781 AVYENADIYLPDDPLSAVDANVGHKIFENVIQPKGMLKNKTRILVITSSMSLPOVDYIIV 840
781 AVYENADIYLPDDPLSAVDANVGHKIFENVIQPKGMLKNKTRILVITSSMSLPOVDYIIV 840
841 MSGGKISEMGSYOELLARDAFAEFLRTYASTEOBDAEENGVTGVSQPKAROMENGM 900
841 MSGGKISEMGSYOELLARDAFAEFLRTYASTEOBDAEENGVTGVSQPKAROMENGM 900
901 LVTSAGKOLOROLSSSSSYSGDISRHNSSTAELOKAKKEEFTWKLMEADKAOTGOYKL 960
901 LVTSAGKOLOROLSSSSSYSGDISRHNSSTAELOKAKKEEFTWKLMEADKAOTGOYKL 960
961 SVYDYMKAIGLFTSFLSIFLMCNHVSALASNYWLSMTDDPIVNGTOEHTKYRLSYVG 1020
961 SVYDYMKAIGLFTSFLSIFLMCNHVSALASNYWLSMTDDPIVNGTOEHTKYRLSYVG 1020

961 SVYDYMKAIGLFTSFLSIFLMCNHVSALASNYWLSMTDDPIVNGTOEHTKYRLSYVG 1020
1021 ALGISOGIANGVGYMAVSIIGIILASRCLHDVLLHSILRSPMSFFERTPSGULVNRFSKEL 1080
1021 ALGISOGIANGVGYMAVSIIGIILASRCLHDVLLHSILRSPMSFFERTPSGULVNRFSKEL 1080
1081 DTVDSMPEVIMKMGSLFENVIGACIVILLATPIAAIIPPLGILYFPQRFYVASSROL 1140
1081 DTVDSMPEVIMKMGSLFENVIGACIVILLATPIAAIIPPLGILYFPQRFYVASSROL 1140
1141 KRLESVSRSPVYSHNETLGVSVIRAFEEQEREIHOSDLKYDENOKAYPSIVANRWLA 1200
1141 KRLESVSRSPVYSHNETLGVSVIRAFEEQEREIHOSDLKYDENOKAYPSIVANRWLA 1200
1201 VRLCEVGNCTVLPALPAVISRHSLSGLVSLSYSLQVTTYANWLVRMSSEMETNIVA 1260
1201 VRLCEVGNCTVLPALPAVISRHSLSGLVSLSYSLQVTTYANWLVRMSSEMETNIVA 1260
1261 VERLKEYSETEKEAPMOIQETAPSSWPQVGRVFRNYCLARYEDLDVLRHINVTINGG 1320
1261 VERLKEYSETEKEAPMOIQETAPSSWPQVGRVFRNYCLARYEDLDVLRHINVTINGG 1320
1321 EKVGI VERTGAGKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
1321 EKVGI VERTGAGKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
1381 SGSLRPMULDPPSQVSDSEBWTSLLELAHLKDFVSLPKLDHECAEGGENLSVGRQOLVCL 1440
1381 SGSLRPMULDPPSQVSDSEBWTSLLELAHLKDFVSLPKLDHECAEGGENLSVGRQOLVCL 1440
1441 ABALLRKTILVDEATAVADLETDDLIQSTIRQFEDCTVLTJAHRAINTMDTRVVL 1500
1441 ABALLRKTILVDEATAVADLETDDLIQSTIRQFEDCTVLTJAHRAINTMDTRVVL 1500
1501 DKGEIOBYGAPSDLLQORGLFYSMAXDAGLVGGGGGMLSRKGIIPBEYVLTSLAEPPAE 1560
1501 DKGEIOBYGAPSDLLQORGLFYSMAXDAGLVGGGGGMLSRKGIIPBEYVLTSLAEPPAE 1560
1561 PRYTRERRARFVSKKNCNVAHNIREQGFLODVFETTLVDLKWPHTLILFTMSFLCSW 1620
1561 PRYTRERRARFVSKKNCNVAHNIREQGFLODVFETTLVDLKWPHTLILFTMSFLCSW 1620
1621 LLFAMVWMLTAPAHG-----DLAP-----GEGTNVPCVTSIHSSSAFLF 1660
1621 LLFAMVWMLTAPAHG-----DLAP-----GEGTNVPCVTSIHSSSAFLF 1660
1661 SIEVOVITIGFGRWTECEPLAILILIVONIVGIMINAMGCIFFMKTAAQHRRAETLIF 1720
1661 SIEVOVITIGFGRWTECEPLAILILIVONIVGIMINAMGCIFFMKTAAQHRRAETLIF 1720
1721 SKHAVITLRHGRLCFMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENG 1780
1721 SKHAVITLRHGRLCFMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENG 1780
1741 SKHAVITLRHGRLCFMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENG 1800
1741 SKHAVITLRHGRLCFMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENG 1800
1781 VGGNGIFLVAPLIIYHYIDNSPLXDLABSDLHHODLEIIVLEGIVETTGITTOARTS 1840
1781 VGGNGIFLVAPLIIYHYIDNSPLXDLABSDLHHODLEIIVLEGIVETTGITTOARTS 1840
1801 VGGNGIFLVAPLIIYHYIDNSPLXDLABSDLHHODLEIIVLEGIVETTGITTOARTS 1860
1801 VGGNGIFLVAPLIIYHYIDNSPLXDLABSDLHHODLEIIVLEGIVETTGITTOARTS 1860
1841 YLADIELMGRFVYVAEEDGRYSVDYSKSGNTIKVTPICTAARQLEDSDS 1891
1841 YLADIELMGRFVYVAEEDGRYSVDYSKSGNTIKVTPICTAARQLEDSDS 1891
1861 YLADIELMGRFVYVAEEDGRYSVDYSKSGNTIKVTPICTAARQLEDSDS 1911
1861 YLADIELMGRFVYVAEEDGRYSVDYSKSGNTIKVTPICTAARQLEDSDS 1911

RESULT 5
US-09-939-853A-86
Sequence 86, Application US/09939853A
Publication No. US20040039163A1
GENERAL INFORMATION:
APPLICANT: Burgess et al.
TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-099
CURRENT APPLICATION NUMBER: US/09/939,853A
CURRENT FILING DATE: 2001-08-27

```

; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 1531
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-939-853A-86

Query Match      80.7%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALGFCGADGSDPLMDNNTWNTSNPDFTKCFONTVLWVPCFYLMACFPFFLYLSRH 60
DB 1 MALGFCGADGSDPLMDNNTWNTSNPDFTKCFONTVLWVPCFYLMACFPFFLYLSRH 60
QY 61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPFLGITTLLA 120
DB 61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPFLGITTLLA 120
QY 121 TFLQLERKRGVQSSGIMLTFWLVALVCAALALRSKIMTALKEQVQLFRDITFYVFS 180
DB 121 TFLQLERKRGVQSSGIMLTFWLVALVCAALALRSKIMTALKEQVQLFRDITFYVFS 180
QY 181 LLLIQLVLSGSDSPLFSETIHDNPPCESASFLSRITFWMIGLIVRGYROPLEGSD 240
DB 181 LLLIQLVLSGSDSPLFSETIHDNPPCESASFLSRITFWMIGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPQRESSKYDANEVEAL 300
DB 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPQRESSKYDANEVEAL 300
QY 301 IVKSPQKMNPSLPKVLVKTGPFLMSFPFKAHDLMMFSGPOLIKLIFKVNDTKAPD 360
DB 301 IVKSPQKMNPSLPKVLVKTGPFLMSFPFKAHDLMMFSGPOLIKLIFKVNDTKAPD 360
QY 361 WQGFYVYTLFVYACLOTLVHLQYFHCFSGMRIKTAVIGAVYKALVITNSAKKSTV 420
DB 361 WQGFYVYTLFVYACLOTLVHLQYFHCFSGMRIKTAVIGAVYKALVITNSAKKSTV 420
QY 421 GEIVNLSVDAQRFDLATYINMISAPLOVITLALYLLMLNIGPSVLAVAVMVLMEVN 480
DB 421 GEIVNLSVDAQRFDLATYINMISAPLOVITLALYLLMLNIGPSVLAVAVMVLMEVN 480
QY 481 AVMAKTKTYOVANHKSNDNRKIMNELINGIKVLKYAMELAFQDKVLAIROBELKYLK 540
DB 481 AVMAKTKTYOVANHKSNDNRKIMNELINGIKVLKYAMELAFQDKVLAIROBELKYLK 540
QY 541 KSAVLSAVGFTWCTPFLVALCFPAVYVTTDENNILLAOQAFVSLALFNILRPPLNLP 600
DB 541 KSAVLSAVGFTWCTPFLVALCFPAVYVTTDENNILLAOQAFVSLALFNILRPPLNLP 600
QY 601 MVISIVQASVSLKRLIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
DB 601 MVISIVQASVSLKRLIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
QY 661 LMGITFSLPEGALVAVGVQVGGKSSLSALLAENDKXEGHVAIKGSVAVYVQOAMIOND 720
DB 661 LMGITFSLPEGALVAVGVQVGGKSSLSALLAENDKXEGHVAIKGSVAVYVQOAMIOND 720
QY 721 SLRENIILGCOLLEBYRSVIOACALPLLEILPSGDRTEIEGKGVNLSGGOKGVSLAR 780
DB 721 SLRENIILGCOLLEBYRSVIOACALPLLEILPSGDRTEIEGKGVNLSGGOKGVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDAAVHGKHFENVIGPKGMLKNKTRILVTHSMSTLPQVDVIV 840
DB 781 AVYSNADIYLFDDPLSAVDAAVHGKHFENVIGPKGMLKNKTRILVTHSMSTLPQVDVIV 840
```

```

DB 781 AVYSNADIYLFDDPLSAVDAAVHGKHFENVIGPKGMLKNKTRILVTHSMSTLPQVDVIV 840
QY 841 MSGGKISEMSGYOEILARRDGAFAEFLRTYASTBOEODAEENGTVGSGPGEAKOMENGM 900
DB 841 MSGGKISEMSGYOEILARRDGAFAEFLRTYASTBOEODAEENGTVGSGPGEAKOMENGM 900
QY 901 LVYDSAGKOLQOROLSSSSSYSGDISRHHSNTAELOKAEKKEETWKLMEADKAOTGVYL 960
DB 901 LVYDSAGKOLQOROLSSSSSYSGDISRHHSNTAELOKAEKKEETWKLMEADKAOTGVYL 960
QY 961 SVYWDYKAIQGLFISFLSIFLPMCNHVSALASNYWLSLWDDPIVNGTOEHTKRLSYG 1020
DB 961 SVYWDYKAIQGLFISFLSIFLPMCNHVSALASNYWLSLWDDPIVNGTOEHTKRLSYG 1020
QY 1021 ALGISOGIAVGYMAVSIIGIILASRCGLHVDLHSHIIRSPMSFEPTPSGULVNRFSKEL 1080
DB 1021 ALGISOGIAVGYMAVSIIGIILASRCGLHVDLHSHIIRSPMSFEPTPSGULVNRFSKEL 1080
QY 1081 DTVDSMIPEVIKMGMSLFNVIGACIYILATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
DB 1081 DTVDSMIPEVIKMGMSLFNVIGACIYILATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
QY 1141 KRLSVSRSPYSHFNELTGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSVSRSPYSHFNELTGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLSCVNGCIVLFAALFAVISRHSLSAGLVLSYSIQVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLSCVNGCIVLFAALFAVISRHSLSAGLVLSYSIQVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIETAPSSPQVGRVEFRNYCLRYRDLDFVLHINVTINGG 1320
DB 1261 VERLKEYSETEKAPWQIETAPSSPQVGRVEFRNYCLRYRDLDFVLHINVTINGG 1320
QY 1321 EKVGVIGRTGAGKSLTGLFRINESAGEIINDGINAKIGLHDLRKKITIIIPQDPVL 1380
DB 1321 EKVGVIGRTGAGKSLTGLFRINESAGEIINDGINAKIGLHDLRKKITIIIPQDPVL 1380
QY 1381 SSSLRMNLDPSSQVSDDEEVTSLSLAHKDVSLAPDLDBECAGGENLSVGOROLVCL 1440
DB 1381 SSSLRMNLDPSSQVSDDEEVTSLSLAHKDVSLAPDLDBECAGGENLSVGOROLVCL 1440
QY 1441 ARALLRKTILVLEATTAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ARALLRKTILVLEATTAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYMAKDAGLV 1531

RESULT 6
US-10-667-891-6
; Sequence 6, Application US/10667891
; Publication No. US20040171024A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BREY, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRAILLES, MARINE
; APPLICANT: RZHETSKY, ANDREY
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
; TITLE OF INVENTION: ANOPHELES
; FILE REFERENCE: 03495-0294-0000
; CURRENT APPLICATION NUMBER: US/10/667, 891
; PRIOR FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: 60/413,469
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 6
; LENGTH: 1531
```

TYPE: PRT
ORGANISM: Homo sapiens
US-10-667-891-6

Query Match 80.7%; Score 7860; DB 4; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRFGSADGSDPLMDMNTWNTSNPDTKCFONTVLVWPCFYLMACPFYFLYSRH 60
DB 1 MALRFGSADGSDPLMDMNTWNTSNPDTKCFONTVLVWPCFYLMACPFYFLYSRH 60

QY 61 DRGVIQMTPLNKTALGFLIMIVCMADLFYSFWERSGFIPLAVFLVSPFLGITTLLA 120
DB 61 DRGVIQMTPLNKTALGFLIMIVCMADLFYSFWERSGFIPLAVFLVSPFLGITTLLA 120

QY 121 TFLIOLERRRKGVOSSGIMLTFMLVALVCAALILRSKIMTALKEDAOYDLFRDITFYVFS 180
DB 121 TFLIOLERRRKGVOSSGIMLTFMLVALVCAALILRSKIMTALKEDAOYDLFRDITFYVFS 180

QY 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFISRTITWMTGLIVRGYRPLGSD 240
DB 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFISRTITWMTGLIVRGYRPLGSD 240

QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKOPVKNVSSKDPAPPKSSSKVDANEVIAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKOPVKNVSSKDPAPPKSSSKVDANEVIAL 300

QY 301 IVKSPQKEMNPSPLEKVLTKTGPYFLMSFFPKAIDHLMFSGPOLIKLILKFNVDTRAPD 360
DB 301 IVKSPQKEMNPSPLEKVLTKTGPYFLMSFFPKAIDHLMFSGPOLIKLILKFNVDTRAPD 360

QY 361 WQGFYTVLLFVTRACLOTVLHQYFHCIFVSGNRKIKAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTVLLFVTRACLOTVLHQYFHCIFVSGNRKIKAVIGAVYRKALVITNSARKSTV 420

QY 421 GEIYNLMSVDAORFMDLATYINMWSAPLOVYTLALYLLMNLGVSVALGAVVWMLMPVN 480
DB 421 GEIYNLMSVDAORFMDLATYINMWSAPLOVYTLALYLLMNLGVSVALGAVVWMLMPVN 480

QY 481 AVNAMKTKTYOVAMHMSKDNRIKLMNEILNGIKVLKYLAMELAFKDKVLAIRBELKVLK 540
DB 481 AVNAMKTKTYOVAMHMSKDNRIKLMNEILNGIKVLKYLAMELAFKDKVLAIRBELKVLK 540

QY 541 KSAVLAVGFTFWCTPFLVALCTFPAYVITIDENNIIDAQTAFLVSLFENILRPPLNLP 600
DB 541 KSAVLAVGFTFWCTPFLVALCTFPAYVITIDENNIIDAQTAFLVSLFENILRPPLNLP 600

QY 601 MWISSIVQASVSLKRLRIFLSHEELBPSIERRPVKDGGGNSITVNAFTFVARSDPT 660
DB 601 MWISSIVQASVSLKRLRIFLSHEELBPSIERRPVKDGGGNSITVNAFTFVARSDPT 660

QY 661 LMGITFSPREGALVAVVGVCCKSSLSALLAEMDVKEGHVAKGSVAVYPOQAWIOND 720
DB 661 LMGITFSPREGALVAVVGVCCKSSLSALLAEMDVKEGHVAKGSVAVYPOQAWIOND 720

QY 721 SLARENILFGQLEPPYRSYIQAICALIPDLIELPSGRTEIGEXVNLSCGQKQVSLAR 780
DB 721 SLARENILFGQLEPPYRSYIQAICALIPDLIELPSGRTEIGEXVNLSCGQKQVSLAR 780

QY 781 AVVSNADIVLPDPLSVDVAHVGHIFENYIGPKMKNTRILLVHSMYSYLPQVDVIIV 840
DB 781 AVVSNADIVLPDPLSVDVAHVGHIFENYIGPKMKNTRILLVHSMYSYLPQVDVIIV 840

QY 841 MSGGKISEMGSYQELLARDAFAFLRTYASTBOBQDAEENGVTGVSQPKKAKOMENGM 900
DB 841 MSGGKISEMGSYQELLARDAFAFLRTYASTBOBQDAEENGVTGVSQPKKAKOMENGM 900

QY 901 LVTNDSAGKQORQULSSSSSYSGDISRRHNSTAELOKAEKKEFWKLMENDKAGTQVYL 960
DB 901 LVTNDSAGKQORQULSSSSSYSGDISRRHNSTAELOKAEKKEFWKLMENDKAGTQVYL 960

QY 961 SVWMDYKAIGLFISFLSIFLFCMCHVSALASNYWLSLMTDDPIVNGTOHTKVRLSVYG 1020

DB 961 SVWMDYKAIGLFISFLSIFLFCMCHVSALASNYWLSLMTDDPIVNGTOHTKVRLSVYG 1020

QY 1021 ALGISQIAVFGSMASVIGIILASRCLHYDLHSILRSWSPFEPFGNLYNRSEKEL 1080
DB 1021 ALGISQIAVFGSMASVIGIILASRCLHYDLHSILRSWSPFEPFGNLYNRSEKEL 1080

QY 1081 DTWDSMIEVYIKMFMGSLFNVIACIVYILATPAALIIIPPLGLITFYVQRFVASSROL 1140
DB 1081 DTWDSMIEVYIKMFMGSLFNVIACIVYILATPAALIIIPPLGLITFYVQRFVASSROL 1140

QY 1141 KRLESVRSFVYSHFNFTLGVSVIRAFEEQERFIHOSDLKVDENQAYPSIYANRWLA 1200
DB 1141 KRLESVRSFVYSHFNFTLGVSVIRAFEEQERFIHOSDLKVDENQAYPSIYANRWLA 1200

QY 1201 VRLCEVNCIVLPALFAVISRHSLSAGLVGSVSVLSQVTTYINMLVRMSSEMETNI VA 1260
DB 1201 VRLCEVNCIVLPALFAVISRHSLSAGLVGSVSVLSQVTTYINMLVRMSSEMETNI VA 1260

QY 1261 VERLKEYSETEKAPMOIORTAPSSWPQVGRVEFRNYCLRYEDLDVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKAPMOIORTAPSSWPQVGRVEFRNYCLRYEDLDVLRHINVTINGG 1320

QY 1321 EKVGIVRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380

QY 1381 SGIIRNMLDPSQYSDBEVWTSLELAHLKDFVSALPKDLHECAEGGENLSVGRQLVCL 1440
DB 1381 SGIIRNMLDPSQYSDBEVWTSLELAHLKDFVSALPKDLHECAEGGENLSVGRQLVCL 1440

QY 1441 ARALLKTKILVUDEATAVDELTDLLIOSTIRTOFEDCTVLIARLANTIMOTRYIVL 1500
DB 1441 ARALLKTKILVUDEATAVDELTDLLIOSTIRTOFEDCTVLIARLANTIMOTRYIVL 1500

QY 1501 DKGEIOBYGAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOBYGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 7
US-10-807-466-6
Sequence 6, Application US/10807466
Publication No. US20040244066A1
GENERAL INFORMATION:
APPLICANT: ROTH, CHARLES W.
APPLICANT: BREY, PAUL T.
APPLICANT: HOLM, INGE
APPLICANT: GRALLES, MARINE
APPLICANT: RZHEVSKY, ANDREY
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
FILE REFERENCE: 03495.0294-01000
CURRENT APPLICATION NUMBER: US/10/807,466
CURRENT FILING DATE: 2004-03-24
PRIOR APPLICATION NUMBER: 10/667,891
PRIOR FILING DATE: 2003-09-23
PRIOR APPLICATION NUMBER: 60/413,469
PRIOR FILING DATE: 2002-09-26
NUMBER OF SEQ ID NOS: 139
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 6
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo sapiens
US-10-807-466-6

Query Match 80.7%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRFGSADGSDPLMDMNTWNTSNPDTKCFONTVLVWPCFYLMACPFYFLYSRH 60

```
Dh 1 MALRGFCSADGSDLDLMDMNTWMTNSNDFTKCFQNTVYLVWPCFYLMACFPFFYLYLSRH 60
Qy 61 DRGIQMTPLNKKTALGFLIMIVCMADLFYSFMRSGRGIPLAVFLVSPFLIGITITLLA 120
Dh 61 DRGIQMTPLNKKTALGFLIMIVCMADLFYSFMRSGRGIPLAVFLVSPFLIGITITLLA 120
Qy 121 TFLIQLERRKGVSSGIMLTFMWALVALCALILRSKIMTALKEDAQVDLFPDITFYYFS 180
Dh 121 TFLIQLERRKGVSSGIMLTFMWALVALCALILRSKIMTALKEDAQVDLFPDITFYYFS 180
Qy 181 LLLIQLVLSGSDSPLESETIHDNPNCPSSASFLSRTFMWITGLIVRGYRQPLEGSD 240
Dh 181 LLLIQLVLSGSDSPLESETIHDNPNCPSSASFLSRTFMWITGLIVRGYRQPLEGSD 240
Qy 241 LMSLNKEDTSEQVVPVLYKMKKCAKTRKQPVKVVYSSKXDAOPKSSKYDANEVEAL 300
Dh 241 LMSLNKEDTSEQVVPVLYKMKKCAKTRKQPVKVVYSSKXDAOPKSSKYDANEVEAL 300
Qy 301 IVKSPOKEMNPSLKVLYKTGPFYFLMSFFPKAIDHLMFSGPOLKLLIFVNDYKAPD 360
Dh 301 IVKSPOKEMNPSLKVLYKTGPFYFLMSFFPKAIDHLMFSGPOLKLLIFVNDYKAPD 360
Qy 361 WQGFYFVTLFVTKCLOTLVHOFHICFVSQWRIKTAIVIGAVYRKALVITNSARKSTV 420
Dh 361 WQGFYFVTLFVTKCLOTLVHOFHICFVSQWRIKTAIVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVILALYLLMLNGPSVLAVGAVMVLMPVN 480
Dh 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVILALYLLMLNGPSVLAVGAVMVLMPVN 480
Qy 481 AVNANKTKTYVAMKSKDNRIKLMNEILNGIKVLYKLYAMBLAFKDKYLAIRQSELKYLK 540
Dh 481 AVNANKTKTYVAMKSKDNRIKLMNEILNGIKVLYKLYAMBLAFKDKYLAIRQSELKYLK 540
Qy 541 KSAVLASVGTFTWCTPFLVALCTPAVYVITDENNIILDAQRAFSALFNLRLPINTLP 600
Dh 541 KSAVLASVGTFTWCTPFLVALCTPAVYVITDENNIILDAQRAFSALFNLRLPINTLP 600
Qy 601 MVISIVQASVSLKRLRIFLSHEELEPDSIERPVKDGGSITVRNATETMARSDPPT 660
Dh 601 MVISIVQASVSLKRLRIFLSHEELEPDSIERPVKDGGSITVRNATETMARSDPPT 660
Qy 661 LMGITFSPREGALVAVVQVCGKSSLSALLAEMDKVEGHVAKGSVAVYPOQAMIOND 720
Dh 661 LMGITFSPREGALVAVVQVCGKSSLSALLAEMDKVEGHVAKGSVAVYPOQAMIOND 720
Qy 721 SLRENILFGCOLERPYSRVIOACALLPDLILPSGDTEIGEGVNLISGQOKQVSLAR 780
Dh 721 SLRENILFGCOLERPYSRVIOACALLPDLILPSGDTEIGEGVNLISGQOKQVSLAR 780
Qy 781 AVYSNADIYLFDDPLSAVDAAVGHKIFENVIQPKGMLKNKTRILVTHSMSTLPQVDVITV 840
Dh 781 AVYSNADIYLFDDPLSAVDAAVGHKIFENVIQPKGMLKNKTRILVTHSMSTLPQVDVITV 840
Qy 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOQDAEENGVTGVSQPKAKOMENG 900
Dh 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOQDAEENGVTGVSQPKAKOMENG 900
Qy 901 LVMTSACKOLOROLSSSSSSSGDTSRHNSSTAELOKAKAKEETWKLMEADKAOQGVKL 960
Dh 901 LVMTSACKOLOROLSSSSSSSGDTSRHNSSTAELOKAKAKEETWKLMEADKAOQGVKL 960
Qy 961 SVYDYDKAKIGLFLISFLSIFLPMCHVSAALASNYWLSMTDDPIVNGIOETHKRLASYG 1020
Dh 961 SVYDYDKAKIGLFLISFLSIFLPMCHVSAALASNYWLSMTDDPIVNGIOETHKRLASYG 1020
Qy 1021 ALGISOGIAVEGYSMAVSIIGILASRCLHVDLHSLHSILSPMSFPERTPSGULVNRFSKEL 1080
Dh 1021 ALGISOGIAVEGYSMAVSIIGILASRCLHVDLHSLHSILSPMSFPERTPSGULVNRFSKEL 1080
Qy 1081 DTUDSMIPREVIKMGMSLFNVIGACTIVILATPIAIIIPPLGLIYFPVOFRYVASSQOL 1140
Dh 1081 DTUDSMIPREVIKMGMSLFNVIGACTIVILATPIAIIIPPLGLIYFPVOFRYVASSQOL 1140
```

```
Qy 1141 KRLESVRSRSPVYSHFNFTLLGVSVIRAFEEQERFIHQSDLKVDENQAVYPSIVANEMLA 1200
Dh 1141 KRLESVRSRSPVYSHFNFTLLGVSVIRAFEEQERFIHQSDLKVDENQAVYPSIVANEMLA 1200
Qy 1201 VRLCEVNCIVLPALFAVAVISRHSLSAGLVGLSVSYLOVTTYINMLVRMSSEMETNIVA 1260
Dh 1201 VRLCEVNCIVLPALFAVAVISRHSLSAGLVGLSVSYLOVTTYINMLVRMSSEMETNIVA 1260
Qy 1261 VERLKEYSERKEAPMOIOETAPPSSWPOYGRVFRNYCLRYREDDLFPVLRHINVTINGG 1320
Dh 1261 VERLKEYSERKEAPMOIOETAPPSSWPOYGRVFRNYCLRYREDDLFPVLRHINVTINGG 1320
Qy 1321 EKVGIVGRTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITTIIPODEVLF 1380
Dh 1321 EKVGIVGRTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITTIIPODEVLF 1380
Qy 1381 SSGIRNMLDPPSOYSDDEWMTSLFLAHLKDFVSALPKLDBHECAGEGNSVGGROLVCL 1440
Dh 1381 SSGIRNMLDPPSOYSDDEWMTSLFLAHLKDFVSALPKLDBHECAGEGNSVGGROLVCL 1440
Qy 1441 ARALLRKTILVLDENATRAAVDLETFDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVIL 1500
Dh 1441 ARALLRKTILVLDENATRAAVDLETFDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVIL 1500
Qy 1501 DKGEIOGYGAPSDLLQQRGLFYSMAKADAGLV 1531
Dh 1501 DKGEIOGYGAPSDLLQQRGLFYSMAKADAGLV 1531

RESULT 8
US-10-484-577-678
; Sequence 678, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft
; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UCT1A1
; FILE REFERENCE: P2285PCT-1
; CURRENT APPLICATION NUMBER: US/10/484,577
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: PCT/EP 02/08220
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: EP 01 11 7608.8
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: EP 02011710.7
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 683
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 678
; LENGTH: 1531
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-484-577-678

Query Match 80.7%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


QY 241 LMSLNKEDTSEOVVPLVVKWKKCECAKTRKOPVKVVYSSKDPAPQKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEOVVPLVVKWKKCECAKTRKOPVKVVYSSKDPAPQKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLKTFGPFYFLMSFFPKAIHDLMMFSGPQILKLLIFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFKVLKTFGPFYFLMSFFPKAIHDLMMFSGPQILKLLIFVNDTKAPD 360
QY 361 WQGFYTYLLFVYTAQLTLVHQYPHICFVSGMRKTKAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYTYLLFVYTAQLTLVHQYPHICFVSGMRKTKAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAORFMDLTYTINMTISAPLOVTLALYTLMLNGSPSVLAGVAVMLMVVN 480
Db 421 GEIVNLMSVDAORFMDLTYTINMTISAPLOVTLALYTLMLNGSPSVLAGVAVMLMVVN 480
QY 481 AVMAKTKTYQVAMHKS KDNRIKLMNEILNGIKVLKLYAMELAFKDYLAIROBELKYLK 540
Db 481 AVMAKTKTYQVAMHKS KDNRIKLMNEILNGIKVLKLYAMELAFKDYLAIROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVTIDENNILDAQTAFVSLAFNILREPILP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVTIDENNILDAQTAFVSLAFNILREPILP 600
QY 601 MVISIVQASVSLKRLRIEFSHEELEPDSTERRPKQCGGCTNSTTVRAATTTWASDPPT 660
Db 601 MVISIVQASVSLKRLRIEFSHEELEPDSTERRPKQCGGCTNSTTVRAATTTWASDPPT 660
QY 661 LMGITFSPGALVAVVGQVCGKSSLSALLAEMDKYGHVAKGSVAAYPOQAMIOND 720
Db 661 LMGITFSPGALVAVVGQVCGKSSLSALLAEMDKYGHVAKGSVAAYPOQAMIOND 720
QY 721 SLRENILFGCOLLEBPYRSVIOACALLPDLILPBGDTEIGERGVNLSSGQKORVSLAR 780
Db 721 SLRENILFGCOLLEBPYRSVIOACALLPDLILPBGDTEIGERGVNLSSGQKORVSLAR 780
QY 781 AVYNAADITYLEDDPLSANDAHVGHIFENVIQPKMLKNKRIIVLTHSMSTLPQVDVIV 840
Db 781 AVYNAADITYLEDDPLSANDAHVGHIFENVIQPKMLKNKRIIVLTHSMSTLPQVDVIV 840
QY 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKOMENGM 900
QY 901 LVTDSAGKQLOQLSSSSSSSGDISRHNSSTAELQKAEKKEETWKLMEADKAQTGYKL 960
Db 901 LVTDSAGKQLOQLSSSSSSSGDISRHNSSTAELQKAEKKEETWKLMEADKAQTGYKL 960
QY 961 SYVYNDYMKALGLFISFLSIFLFCMCHVSALASNTWLSMTDDPIVNGTOETKXRLSYG 1020
Db 961 SYVYNDYMKALGLFISFLSIFLFCMCHVSALASNTWLSMTDDPIVNGTOETKXRLSYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFFERTPSGULVNFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFFERTPSGULVNFSEKL 1080
QY 1081 DTVDSMTPEVILKMGSLFENVIGACIVILLATPIAIIIPPLGLIYFVQGFYASSROL 1140
Db 1081 DTVDSMTPEVILKMGSLFENVIGACIVILLATPIAIIIPPLGLIYFVQGFYASSROL 1140
QY 1141 KRLESVRSRPSYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIYANRWLA 1200
Db 1141 KRLESVRSRPSYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIYANRWLA 1200
QY 1201 VRLECVNCTIVLFALFAVISRHSLSAGLVGLSVYSLOVTTYLNMLVRMSSMETNIVA 1260
Db 1201 VRLECVNCTIVLFALFAVISRHSLSAGLVGLSVYSLOVTTYLNMLVRMSSMETNIVA 1260
QY 1261 VERLKEYSETKEAPMOIQETAPSSWPQVGRVFEFRNYCLATYREDLDFVLRIHINVTTNGC 1320
Db 1261 VERLKEYSETKEAPMOIQETAPSSWPQVGRVFEFRNYCLATYREDLDFVLRIHINVTTNGC 1320

QY 1321 EKVGIAGRTGAGKSSLTGLFRINESAGEIILIDGINAKIGLHDLRFKTIIPODPVLF 1380
Db 1321 EKVGIAGRTGAGKSSLTGLFRINESAGEIILIDGINAKIGLHDLRFKTIIPODPVLF 1380
QY 1381 SGLRMLNDPFSQYSDEEWTSLELAHLKDFVSALPDGLDHECAEGENLSVGROLYCL 1440
Db 1381 SGLRMLNDPFSQYSDEEWTSLELAHLKDFVSALPDGLDHECAEGENLSVGROLYCL 1440
QY 1441 ARALLRTKILVLEDAITPAVDLEFDDLIQSTIRIROPEDCTVLTAAHRLNTIMDTIRYVL 1500
Db 1441 ARALLRTKILVLEDAITPAVDLEFDDLIQSTIRIROPEDCTVLTAAHRLNTIMDTIRYVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMAKXAGIV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYSMAKXAGIV 1531

RESULT 9
US-10-889-503-19
; Sequence 19, Application US/10889503
; Publication No. US20050063968A1
; GENERAL INFORMATION:
; APPLICANT: Fox Chase Cancer Center
; APPLICANT: Krüh, Gary D.
; APPLICANT: Lee, Kun
; APPLICANT: Belinsky, Martin G.
; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
; FILE REFERENCE: FCCC 98-02
; CURRENT APPLICATION NUMBER: US/10/889, 503
; PRIOR FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: US/09/647, 140
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US99/06644
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/079, 759
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/095, 153
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-889-503-19

Query Match 80.7%; Score 7860; DB 5; Length 1531;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCAGSDPLDMNNTANTSNPDKCFONTLVWVPCYLMACPPFVLYLSRH 60
Db 1 MALRGFCAGSDPLDMNNTANTSNPDKCFONTLVWVPCYLMACPPFVLYLSRH 60
QY 61 DRGYIQTPLNKTKTALGFLIMIVCWADLFYSEWERSRGIFLAPVFLVSPTLGITTLLA 120
Db 61 DRGYIQTPLNKTKTALGFLIMIVCWADLFYSEWERSRGIFLAPVFLVSPTLGITTLLA 120
QY 121 TPLIQLERRRGVSSGIMLTFWVALVCAALILRSKIMTLAKEDAOVDLFRDITFYVYS 180
Db 121 TPLIQLERRRGVSSGIMLTFWVALVCAALILRSKIMTLAKEDAOVDLFRDITFYVYS 180
QY 181 LLLIQLVLSGFSRPSPLFSETIHDPNCPRESSASFLSRITFMMITGLIVGYRPLBSPD 240
Db 181 LLLIQLVLSGFSRPSPLFSETIHDPNCPRESSASFLSRITFMMITGLIVGYRPLBSPD 240
QY 241 LMSLNKEDTSEOVVPLVVKWKKCECAKTRKOPVKVVYSSKDPAPQKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEOVVPLVVKWKKCECAKTRKOPVKVVYSSKDPAPQKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLKTFGPFYFLMSFFPKAIHDLMMFSGPQILKLLIFVNDTKAPD 360

Db 301 IVKSPQKEMNSLFRVLYKTEGPFYFLMSFFPKAHDLMFSGPOLIKLIFVNDTKAPD 360
Qy 361 WQGYFYTVLLFVTACLOTLVTHQYFHI CFVSGMR IKTA VIGAVYKALVITNSARKSSTV 420
Db 361 WQGYFYTVLLFVTACLOTLVTHQYFHI CFVSGMR IKTA VIGAVYKALVITNSARKSSTV 420
Qy 421 GEIVNLMSVDAQRFMDLATYINMTWSAPLOVILALYLLMLNGPSVLGAVAVMLMVEVN 480
Db 421 GEIVNLMSVDAQRFMDLATYINMTWSAPLOVILALYLLMLNGPSVLGAVAVMLMVEVN 480
Qy 481 AVMAAMKTTYQVAHMKSKDNRIKLMNEILNGIKYVLA YAMELAFKDKYLAIRQEBLKYLK 540
Db 481 AVMAAMKTTYQVAHMKSKDNRIKLMNEILNGIKYVLA YAMELAFKDKYLAIRQEBLKYLK 540
Qy 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQTA FVSALFNILRPPLNTLP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQTA FVSALFNILRPPLNTLP 600
Qy 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGCTNSITVRNATFTWASDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGCTNSITVRNATFTWASDPT 660
Qy 661 LMGITFSPREGALVAVVQVQGGKSSLSALLAENDKXEGHVAIKGSVAVYPOQAMIOND 720
Db 661 LMGITFSPREGALVAVVQVQGGKSSLSALLAENDKXEGHVAIKGSVAVYPOQAMIOND 720
Qy 721 SLRENIIFGCOLBEEYYRSVIOACALPDLLEILPBGSDTEIGEKVNLSSGOKOVSLAR 780
Db 721 SLRENIIFGCOLBEEYYRSVIOACALPDLLEILPBGSDTEIGEKVNLSSGOKOVSLAR 780
Qy 781 AVYSNADITLFPDPLSAVDAAHVGKHI FENVIGPKMLXNKRLILVTHSMYLPQVDVILV 840
Db 781 AVYSNADITLFPDPLSAVDAAHVGKHI FENVIGPKMLXNKRLILVTHSMYLPQVDVILV 840
Qy 841 MSGGKISMGSGYOELIARDGAFAEFLRTYASTEQDABENGCVTVSGGKAKOMENGM 900
Db 841 MSGGKISMGSGYOELIARDGAFAEFLRTYASTEQDABENGCVTVSGGKAKOMENGM 900
Qy 901 LVTDSAGOLORUSSSSSGDISRHHNSTAELEKAEKKEETKMLWEADAOQGOYKL 960
Db 901 LVTDSAGOLORUSSSSSGDISRHHNSTAELEKAEKKEETKMLWEADAOQGOYKL 960
Qy 961 SVYMDYMKALGLFLISFLSIFLPMCNHVSALASNYMLSLMTDPIVNGTOEHTKVALSYVG 1020
Db 961 SVYMDYMKALGLFLISFLSIFLPMCNHVSALASNYMLSLMTDPIVNGTOEHTKVALSYVG 1020
Qy 1021 ALGISQGIADVGYSAVSIIGITLARSCLHVDLHSLNSPMSFFERTPSGNLVNRFSKEL 1080
Db 1021 ALGISQGIADVGYSAVSIIGITLARSCLHVDLHSLNSPMSFFERTPSGNLVNRFSKEL 1080
Qy 1081 DTVDSMIPEVIMKMFNGSLFNVIGACIVILATPILAIITIPGLLYFFQRFYVASSRQL 1140
Db 1081 DTVDSMIPEVIMKMFNGSLFNVIGACIVILATPILAIITIPGLLYFFQRFYVASSRQL 1140
Qy 1141 KRLESVSHSPVYSHNETLLAGSVIRAFEEORFIHQSDLKVDENOKAYPSIVANRWLA 1200
Db 1141 KRLESVSHSPVYSHNETLLAGSVIRAFEEORFIHQSDLKVDENOKAYPSIVANRWLA 1200
Qy 1201 VRLBECVNCIYVPAALPAVISHSLSAGLVGSYSISLOVTTYLNLVMSSEMETNIVA 1260
Db 1201 VRLBECVNCIYVPAALPAVISHSLSAGLVGSYSISLOVTTYLNLVMSSEMETNIVA 1260
Qy 1261 VERLKEYETBEKEAWOIOETAPRSSWPOVGRVPRPNYCLARTREDLDFLRHINVTINGG 1320
Db 1261 VERLKEYETBEKEAWOIOETAPRSSWPOVGRVPRPNYCLARTREDLDFLRHINVTINGG 1320
Qy 1321 EKVGIVGRTAGKSSSLTGLFRINSAGEEIIIDGINAKIGLHDLRFKITIIPODPVLF 1380
Db 1321 EKVGIVGRTAGKSSSLTGLFRINSAGEEIIIDGINAKIGLHDLRFKITIIPODPVLF 1380
Qy 1381 SGSLRMLNIDPFSQYSDEBVTLSLELAHLKD FVSALPDKLDHBCAEGENLSVGORQVCL 1440
Db 1381 SGSLRMLNIDPFSQYSDEBVTLSLELAHLKD FVSALPDKLDHBCAEGENLSVGORQVCL 1440

Db 1381 SGSLRMLNIDPFSQYSDEBVTLSLELAHLKD FVSALPDKLDHBCAEGENLSVGORQVCL 1440
Qy 1441 ARALIRKTKIILVUDEATAVADLETDLIQSTIRTOFEDCTVLIHARLNTIMDYTRVIL 1500
Db 1441 ARALIRKTKIILVUDEATAVADLETDLIQSTIRTOFEDCTVLIHARLNTIMDYTRVIL 1500
Qy 1501 DKGEIOBYGAPSDLLQORGLFYSMAKONAGLV 1531
Db 1501 DKGEIOBYGAPSDLLQORGLFYSMAKONAGLV 1531

RESULT 10
US-10-756-149-5033
; Sequence 5033, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Naeem
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756.149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 5033
; LENGTH: 1531
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-756-149-5033

Query Match 80.7%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALGFGCADSDPLMDNVTWNTSNPFTGCFQYTVLVWVPCRYLWACFPFYLYLSRH 60
Db 1 MALGFGCADSDPLMDNVTWNTSNPFTGCFQYTVLVWVPCRYLWACFPFYLYLSRH 60
Qy 61 DRGYIOWPLNKTKTALGFLMTICMADLFYSPMERSGIFLAPFVLSPTLIGITTLA 120
Db 61 DRGYIOWPLNKTKTALGFLMTICMADLFYSPMERSGIFLAPFVLSPTLIGITTLA 120
Qy 121 TFLIQLERRKVOSSGIMLTFMLVALYCALAILRSKIMTALKEBAQVDLPFDITFYVFS 180
Db 121 TFLIQLERRKVOSSGIMLTFMLVALYCALAILRSKIMTALKEBAQVDLPFDITFYVFS 180
Qy 181 LLLIOLVLSGFSDBSPLESETIHPNCPRESSASFLSITTTWMTTGLIVRGYRQPLEGSD 240
Db 181 LLLIOLVLSGFSDBSPLESETIHPNCPRESSASFLSITTTWMTTGLIVRGYRQPLEGSD 240
Qy 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKOPVKVYSSXDPAPOKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKOPVKVYSSXDPAPOKESSKYDANEVEAL 300
Qy 301 IVKSPQKEMNSLFRVLYKTEGPFYFLMSFFPKAHDLMFSGPOLIKLIFVNDTKAPD 360
Db 301 IVKSPQKEMNSLFRVLYKTEGPFYFLMSFFPKAHDLMFSGPOLIKLIFVNDTKAPD 360
Qy 361 WQGYFYTVLLFVTACLOTLVTHQYFHI CFVSGMR IKTA VIGAVYKALVITNSARKSSTV 420
Db 361 WQGYFYTVLLFVTACLOTLVTHQYFHI CFVSGMR IKTA VIGAVYKALVITNSARKSSTV 420
Qy 421 GEIVNLMSVDAQRFMDLATYINMTWSAPLOVILALYLLMLNGPSVLGAVAVMLMVEVN 480
Db 421 GEIVNLMSVDAQRFMDLATYINMTWSAPLOVILALYLLMLNGPSVLGAVAVMLMVEVN 480
Qy 481 AVMAAMKTTYQVAHMKSKDNRIKLMNEILNGIKYVLA YAMELAFKDKYLAIRQEBLKYLK 540
Db 481 AVMAAMKTTYQVAHMKSKDNRIKLMNEILNGIKYVLA YAMELAFKDKYLAIRQEBLKYLK 540
Qy 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQTA FVSALFNILRPPLNTLP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQTA FVSALFNILRPPLNTLP 600

Db 541 KSAVLSAVGTFWVCTPFLVALCTFAVYVTIDENNILDAQTAFAVSLALFNILRPPLNLP 600
 Qy 601 MVSISYQASVSLRLRIFLSHEELPDSIERRPVKGGGNSITVNAATPTARSPPT 660
 Db 601 MVSISYQASVSLRLRIFLSHEELPDSIERRPVKGGGNSITVNAATPTARSPPT 660
 Qy 661 LMGITFSIPGALVAVVGVCGGSSLSALLAEMDKVGHVAIKGSVAVYPOQAMQND 720
 Db 661 LMGITFSIPGALVAVVGVCGGSSLSALLAEMDKVGHVAIKGSVAVYPOQAMQND 720
 Qy 721 SLRNNILFGCOLPEPYSVIAQACALLPDLILPSGDRTEIGEKVNLGGQKQVSLAR 780
 Db 721 SLRNNILFGCOLPEPYSVIAQACALLPDLILPSGDRTEIGEKVNLGGQKQVSLAR 780
 Qy 781 AVYSNADIVLPDDLAVDAVGHKIFENVIKPGKMLKNTKRIIVTSHMSLTPVDYIV 840
 Db 781 AVYSNADIVLPDDLAVDAVGHKIFENVIKPGKMLKNTKRIIVTSHMSLTPVDYIV 840
 Qy 841 MSGGKISEMGSYOELARDGAFAEFLRTYASTBOEODAEENGTVGSGPKEAKQOMENG 900
 Db 841 MSGGKISEMGSYOELARDGAFAEFLRTYASTBOEODAEENGTVGSGPKEAKQOMENG 900
 Qy 901 LYTDSAGKOLQROLSSSSSSYSGDISRRHNSTAELOKAEKKEETWKLMEADKAQTQVKL 960
 Db 901 LYTDSAGKOLQROLSSSSSSYSGDISRRHNSTAELOKAEKKEETWKLMEADKAQTQVKL 960
 Qy 961 SYVWDYKAKGLPISFLSTFLPMCNHVSALASNTWLSMTDDPVLNGTOHTKRLSVYG 1020
 Db 961 SYVWDYKAKGLPISFLSTFLPMCNHVSALASNTWLSMTDDPVLNGTOHTKRLSVYG 1020
 Qy 1021 ALGISOGIAVFGYSMAVSIGIILASRCIAVDLHSLIRSPMSFEPRPSGKLVNRFSEKL 1080
 Db 1021 ALGISOGIAVFGYSMAVSIGIILASRCIAVDLHSLIRSPMSFEPRPSGKLVNRFSEKL 1080
 Qy 1081 DTVDSMIPBEVYIKPMFMSLFPVNGACIYLATPILAIIPPLGLIYFVQRFYVASSROL 1140
 Db 1081 DTVDSMIPBEVYIKPMFMSLFPVNGACIYLATPILAIIPPLGLIYFVQRFYVASSROL 1140
 Qy 1141 KRLSVRSRPFYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANFWLA 1200
 Db 1141 KRLSVRSRPFYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANFWLA 1200
 Qy 1201 VRLKESVRSRPFYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANFWLA 1260
 Db 1201 VRLKESVRSRPFYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANFWLA 1260
 Qy 1261 VERLKEYSETEKAPMOIOTETAPSSWPQYGRVFRNYCLRYEDDLFVLRLHNVITNGG 1320
 Db 1261 VERLKEYSETEKAPMOIOTETAPSSWPQYGRVFRNYCLRYEDDLFVLRLHNVITNGG 1320
 Qy 1321 EKVGIYGRTAGAGSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTTIIPQDVLV 1380
 Db 1321 EKVGIYGRTAGAGSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTTIIPQDVLV 1380
 Qy 1381 SGLSRNMLDPFSOYSDSEWWTSLBLAHLKDFVSLPRKLDHECAGEGENSVQORQVCL 1440
 Db 1381 SGLSRNMLDPFSOYSDSEWWTSLBLAHLKDFVSLPRKLDHECAGEGENSVQORQVCL 1440
 Qy 1441 ARALLKRTKILVDEATAVDLETDLIQSTIRTOFEDCTVLTARHLANTIMDYTRYIVL 1500
 Db 1441 ARALLKRTKILVDEATAVDLETDLIQSTIRTOFEDCTVLTARHLANTIMDYTRYIVL 1500
 Qy 1501 DKGEIOEYCAPSDLLQORGLFYSMADAGLIV 1531
 Db 1501 DKGEIOEYCAPSDLLQORGLFYSMADAGLIV 1531

; TITLE OF INVENTION: No. US20040039163a1el Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-099
 ; CURRENT APPLICATION NUMBER: US/09/939,853A
 ; CURRENT FILING DATE: 2001-08-27
 ; PRIOR APPLICATION NUMBER: 60/228,191
 ; PRIOR FILING DATE: 2000-08-25
 ; PRIOR APPLICATION NUMBER: 60/267,300
 ; PRIOR FILING DATE: 2001-02-08
 ; PRIOR APPLICATION NUMBER: 60/269,961
 ; PRIOR FILING DATE: 2001-02-20
 ; PRIOR APPLICATION NUMBER: 60/277,337
 ; PRIOR FILING DATE: 2001-03-20
 ; NUMBER OF SEQ ID NOS: 159
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 87
 ; LENGTH: 1515
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-939-853A-87
 Query Match 79.8%; Score 7769; DB 3; Length 1515;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 17 DNVNTWNTSNPDFTKCFONTLVWVPCFYLMACPPEFLYSRDRGVIQMTPLNKTETA 76
 Db 1 DNVNTWNTSNPDFTKCFONTLVWVPCFYLMACPPEFLYSRDRGVIQMTPLNKTETA 60
 Qy 77 LGFLMLVVCADLDFYSWEWSRGI FLAPVLVSPFTLLGFTLLATFLIQERRKGVOSG 136
 Db 61 LGFLMLVVCADLDFYSWEWSRGI FLAPVLVSPFTLLGFTLLATFLIQERRKGVOSG 120
 Qy 137 IMLTFWLVALVCAIILRSKIMTALXEDAOVDLFRDITFYVYVSLLIQVLSCFSRSP 196
 Db 121 IMLTFWLVALVCAIILRSKIMTALXEDAOVDLFRDITFYVYVSLLIQVLSCFSRSP 180
 Qy 197 LPEETHDPNCPRESSASFLSRITFWITGLIYRGROPLEGDLNLEEDTSEQVVPV 256
 Db 181 LPEETHDPNCPRESSASFLSRITFWITGLIYRGROPLEGDLNLEEDTSEQVVPV 240
 Qy 257 LVNKMKECKATKOPKPVVYSSKDPAPQKSSKVDANEVVALIYKSPKEMNPSLFKV 316
 Db 241 LVNKMKECKATKOPKPVVYSSKDPAPQKSSKVDANEVVALIYKSPKEMNPSLFKV 300
 Qy 317 LYTFGPYPLMSEFPFAIHDMWFSGPQIILKLIKFNVDTKADWQGYFTVLLFVTACL 376
 Db 301 LYTFGPYPLMSEFPFAIHDMWFSGPQIILKLIKFNVDTKADWQGYFTVLLFVTACL 360
 Qy 377 QTLVLMQYFHICVSGMRITVAVYRQALVITNSARKSSVGEIVNIMSDAQRFMD 436
 Db 361 QTLVLMQYFHICVSGMRITVAVYRQALVITNSARKSSVGEIVNIMSDAQRFMD 420
 Qy 437 LATYINMWSAPQVILALYILMLNGPSVLAVAVVWLVNVAWAMTKTYQVVAHMK 496
 Db 421 LATYINMWSAPQVILALYILMLNGPSVLAVAVVWLVNVAWAMTKTYQVVAHMK 480
 Qy 497 SKDNRIKLMEILNGIKVILKYAMELAFKQKVALIROBELKVLKKSAYLSAVGTFTWCT 556
 Db 481 SKDNRIKLMEILNGIKVILKYAMELAFKQKVALIROBELKVLKKSAYLSAVGTFTWCT 540
 Qy 557 PFLVALCTPRAVYTTIDENNILDAQTAFAVSLALFNILRPPLNLPMTYISIVQASVSLKRL 616
 Db 541 PFLVALCTPRAVYTTIDENNILDAQTAFAVSLALFNILRPPLNLPMTYISIVQASVSLKRL 600
 Qy 617 RIFLSHEELPDSIERRPVKGGGNSITVNAATPTARSDPPTLNGITFSIPGALVAV 676
 Db 601 RIFLSHEELPDSIERRPVKGGGNSITVNAATPTARSDPPTLNGITFSIPGALVAV 660
 Qy 677 VGVGCGGSSLSALLAEMDKVGHVAIKGSVAVYPOQAMQNDLSRENILFGCOLPEEPY 736
 Db 661 VGVGCGGSSLSALLAEMDKVGHVAIKGSVAVYPOQAMQNDLSRENILFGCOLPEEPY 720
 Qy 737 YRSVIAQACALLPDLILPSGDRTEIGEKVNLGGQKQVSLARAVYSNADIVLPDDLPS 796

RESULT 11
 US-09-939-853A-87
 ; Sequence 87, Application US/09939853A
 ; Publication No. US20040039163a1
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgees et al.

|||||
Db 721 YRSVIQACALLPDEIILPSGDRTEIGEGKVNLSGGQKORVSLARAVVSNADIIYFDPLS 780
QY 797 AVDAHVGKHIPEENYIGPKGMLKNKTRILIVTHSMSTLPPOVDYIIVMSGGKISEMSYQELI 856
Db 781 AVDAHVGKHIPEENYIGPKGMLKNKTRILIVTHSMSTLPPOVDYIIVMSGGKISEMSYQELI 840
QY 857 ARDGAFAEFLRTYASTBEOQDAEENGVTGSGPGKEAKOMENGMVLTDSAGKQOLROLS 916
Db 841 ARDGAFAEFLRTYASTBEOQDAEENGVTGSGPGKEAKOMENGMVLTDSAGKQOLROLS 900
QY 917 SSSYSGDISRRHNSTAEIQKAEAKKEETWKLMKADKATQGVKLSVYVDYKAIIGLFIISF 976
Db 901 SSSYSGDISRRHNSTAEIQKAEAKKEETWKLMKADKATQGVKLSVYVDYKAIIGLFIISF 960
QY 977 LSIPLFMCHNVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYGALGISOGIAVFGYSMA 1036
Db 961 LSIPLFMCHNVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYGALGISOGIAVFGYSMA 1020
QY 1037 VSIIGIILASRCLAHVDLHSLIRSPMSPEERTPSGNLVNRFSEKLDVTVDMSIPEVIXKMEWG 1096
Db 1021 VSIIGIILASRCLAHVDLHSLIRSPMSPEERTPSGNLVNRFSEKLDVTVDMSIPEVIXKMEWG 1080
QY 1097 SLEFVIGACIYIILATPIAIIIPPLGIIFPVQAFYVASSROLKLESVSRSPIYSHFN 1156
Db 1081 SLEFVIGACIYIILATPIAIIIPPLGIIFPVQAFYVASSROLKLESVSRSPIYSHFN 1140
QY 1157 ETLGVSIVYIRAFEBEORFIHOSDLKVDENOKAYPSIYANRWLAVALCEVGNCIYLPAL 1216
Db 1141 ETLGVSIVYIRAFEBEORFIHOSDLKVDENOKAYPSIYANRWLAVALCEVGNCIYLPAL 1200
QY 1217 PAVISRHSLSAGLVGLSVYSLSQVTTYLNMVLRMSSEMETNIIVAVERLKEYSETEKAPW 1276
Db 1201 PAVISRHSLSAGLVGLSVYSLSQVTTYLNMVLRMSSEMETNIIVAVERLKEYSETEKAPW 1260
QY 1277 QIOETAPSSMPQVGRVFRNYCLARYRDLDPVLRHINVTINGEKRGIVGRTGKSSL 1336
Db 1261 QIOETAPSSMPQVGRVFRNYCLARYRDLDPVLRHINVTINGEKRGIVGRTGKSSL 1320
QY 1337 TLGFRIIESAGEIIDIIGINIAKIGLHDLRPKTIIPQDPVLFSGSLRNNLDPPSOYSD 1396
Db 1321 TLGFRIIESAGEIIDIIGINIAKIGLHDLRPKTIIPQDPVLFSGSLRNNLDPPSOYSD 1380
QY 1397 EEVWTSLELAHKDFVSALPKLDHECAEGGENLSVGOROLVCLARALLRKTKILVLDPA 1456
Db 1381 EEVWTSLELAHKDFVSALPKLDHECAEGGENLSVGOROLVCLARALLRKTKILVLDPA 1440
QY 1457 TAAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILO 1516
Db 1441 TAAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILO 1500
QY 1517 ORGLFYSAKXDAGLV 1531
Db 1501 ORGLFYSAKXDAGLV 1515

RESULT 12
US-10-618-281-42
; Sequence 42, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously Known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-281-42
Query Match 79.8%; Score 7763; DB 5; Length 1515;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 17 DNANTNTSNDPFKCFONTVLVWVPCFYLMACPPFFELYSRRDRYIOWTPIINKTKTA 76
Db 1 DNANTNTSNDPFKCFONTVLVWVPCFYLMACPPFFELYSRRDRYIOWTPIINKTKTA 60
QY 77 LGFLIMIVCANADLFYSWERSRGIFLAPVFLVSPTELLGITLTLATFLIQLERRKGVSSG 136
Db 61 LGFLIMIVCANADLFYSWERSRGIFLAPVFLVSPTELLGITLTLATFLIQLERRKGVSSG 120
QY 137 IMTLFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFSLLLIQVLSCFSDRSP 196
Db 121 IMTLFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFSLLLIQVLSCFSDRSP 180
QY 197 LFSETIHDPNCPSSSASFLSRITFWMITGLIYRGYQPLEGSDLMSLNKEDTSEQVVPV 256
Db 181 LFSETIHDPNCPSSSASFLSRITFWMITGLIYRGYQPLEGSDLMSLNKEDTSEQVVPV 240
QY 257 LVKMKKECAKTRQOPKPVVYSSKDPAQPKSSKYDANEEVEALIVKSPQKRNPSLFKV 316
Db 241 LVKMKKECAKTRQOPKPVVYSSKDPAQPKSSKYDANEEVEALIVKSPQKRNPSLFKV 300
QY 317 LYKTFGPFLMSFPFKAIHDLIMFSGPOLIKLIRFVNDIYAPDWGYFTVLLFTVACL 376
Db 301 LYKTFGPFLMSFPFKAIHDLIMFSGPOLIKLIRFVNDIYAPDWGYFTVLLFTVACL 360
QY 377 QTLVILHQYFHCIFNSGRIRKTAIVGAYRKALVITNSARKSTYGEIVNLSVDAQRMD 436
Db 361 QTLVILHQYFHCIFNSGRIRKTAIVGAYRKALVITNSARKSTYGEIVNLSVDAQRMD 420
QY 437 LATYNNMWSAPLOVIALYLLMLNLPSTVLAGAVWVLMVPAVNAVMKRTKYQVAHMK 496
Db 421 LATYNNMWSAPLOVIALYLLMLNLPSTVLAGAVWVLMVPAVNAVMKRTKYQVAHMK 480
QY 497 SKDNRIKIMNEILINGIKVLYAMELAFKDVLAIRQELKVLKKSAYLSAVGFTWCT 556
Db 481 SKDNRIKIMNEILINGIKVLYAMELAFKDVLAIRQELKVLKKSAYLSAVGFTWCT 540
QY 557 PFLVALCTFAYVYVITDENNIIDAOQAFVSALPNIIRPPLNILEMVTSSIVQASVSLKRL 616
Db 541 PFLVALCTFAYVYVITDENNIIDAOQAFVSALPNIIRPPLNILEMVTSSIVQASVSLKRL 600
QY 617 RIFLSHEELBPSLIERPVKDGGSNTSYVNAATFTWARSDPPTLNGITFSPICGALYAV 676
Db 601 RIFLSHEELBPSLIERPVKDGGSNTSYVNAATFTWARSDPPTLNGITFSPICGALYAV 660
QY 677 VGOVCGCKSSLSALLAEMDVKEGVHAIKGSVAVVPQAWIONDLSRENILFGCQLEBPY 736
Db 661 VGOVCGCKSSLSALLAEMDVKEGVHAIKGSVAVVPQAWIONDLSRENILFGCQLEBPY 720
QY 737 YRSVIQACALLPDEIILPSGDRTEIGEGKVNLSGGQKORVSLARAVVSNADIIYFDPLS 796
Db 721 YRSVIQACALLPDEIILPSGDRTEIGEGKVNLSGGQKORVSLARAVVSNADIIYFDPLS 780
QY 797 AVDAHVGKHIPEENYIGPKGMLKNKTRILIVTHSMSTLPPOVDYIIVMSGGKISEMSYQELI 856
Db 781 AVDAHVGKHIPEENYIGPKGMLKNKTRILIVTHSMSTLPPOVDYIIVMSGGKISEMSYQELI 840
QY 857 ARDGAFAEFLRTYASTBEOQDAEENGVTGSGPGKEAKOMENGMVLTDSAGKQOLROLS 916
Db 841 ARDGAFAEFLRTYASTBEOQDAEENGVTGSGPGKEAKOMENGMVLTDSAGKQOLROLS 900
QY 917 SSSYSGDISRRHNSTAEIQKAEAKKEETWKLMKADKATQGVKLSVYVDYKAIIGLFIISF 976
Db 901 SSSYSGDISRRHNSTAEIQKAEAKKEETWKLMKADKATQGVKLSVYVDYKAIIGLFIISF 960

Db 901 SSSYSGDISRHHNSTAELOKAEKKEETMKLEADKQGVKLSVYDYMKAIGLFI SF 960
Qy 977 LSI FLFMCNHYASALASNYWLSLMTDDPIVNGOEHTKYRLSVYGA LGISOGIAVPGYMA 1036
Db 961 LSI FLFMCNHYASALASNYWLSLMTDDPIVNGOEHTKYRLSVYGA LGISOGIAVPGYMA 1020
Qy 1037 VSI GGIASRC LHVLDLHLSIRSPMSPEPTPSGNLVNRFSEKLDVTYSMTPEVIMKMG 1096
Db 1021 VSI GGIASRC LHVLDLHLSIRSPMSPEPTPSGNLVNRFSEKLDVTYSMTPEVIMKMG 1080
Qy 1097 SLE NVIGACIYI LLA TP IAA I I I P P L G L I Y F V O R F Y A S S R O L K R L E S V S R S P Y S H N 1156
Db 1081 SLE NVIGACIYI LLA TP IAA I I I P P L G L I Y F V O R F Y A S S R O L K R L E S V S R S P Y S H N 1140
Qy 1157 ETLGVSYIRAFEEOEERFIHOSDLKVDENOKAYYPSIVANRWLA VRLECGNCTVLPAL 1216
Db 1141 ETLGVSYIRAFEEOEERFIHOSDLKVDENOKAYYPSIVANRWLA VRLECGNCTVLPAL 1200
Qy 1217 FAVISRHSLSAGLVGLSVYSGLQVTTYLNMLV RMSSEMETNI VAVERLKEYS ETEKEAPW 1276
Db 1201 FAVISRHSLSAGLVGLSVYSGLQVTTYLNMLV RMSSEMETNI VAVERLKEYS ETEKEAPW 1260
Qy 1277 QIQTAPSSWPQYGRVFEFRNYCLRYREDLPVL RHINVTINGEKVGI VGR TGAGKSSL 1336
Db 1261 QIQTAPSSWPQYGRVFEFRNYCLRYREDLPVL RHINVTINGEKVGI VGR TGAGKSSL 1320
Qy 1337 TLGFRINESAEGEIIIDGNI AKI GLHDLAFKTTI IQDBV L FSGSIRMLNDPFSQYSD 1396
Db 1321 TLGFRINESAEGEIIIDGNI AKI GLHDLAFKTTI IQDBV L FSGSIRMLNDPFSQYSD 1380
Qy 1397 EEWTSLELAHLKDFVSALPKLDHECAGGENTSVGOROLVCLARALLRKTILYVDEA 1456
Db 1381 EEWTSLELAHLKDFVSALPKLDHECAGGENTSVGOROLVCLARALLRKTILYVDEA 1440
Qy 1457 TAAVLDLTDLLIOSTIRTOFEDCTVLIARLNTIMDYTRYVL DKGEIOEYGA PSDLQ 1516
Db 1441 TAAVLDLTDLLIOSTIRTOFEDCTVLIARLNTIMDYTRYVL DKGEIOEYGA PSDLQ 1500
Qy 1517 ORGLFYSMAVDAGLV 1531
Db 1501 ORGLFYSMAVDAGLV 1515

RESULT 13
US-10-408-765A-1718
; Sequence 1718, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Baby, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408.765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1718
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1718

Query Match 72.3%; Score 7037; DB 4; Length 1388;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;
Qy 67 MTPINKTKTALGFLMTIVCWADLFYSPWERSRGIFLAPVFLVSPTLGITTLLATFLIQL 126

Db 1 MTPINKTKTALGFLMTIVCWADLFYSPWERSRGIFLAPVFLVSPTLGITTLLATFLIQL 60
Qy 127 ERRKGVSSGIMLTFMIVALYCALAIIIRSKITMALKEDAQVDLFRDITFYVPSLLIQL 186
Db 61 ERRKGVSSGIMLTFMIVALYCALAIIIRSKITMALKE----- 97
Qy 187 VLSCFSRSPLFSETIHDPNCPRESSASFLSRIFMWTGLIVRGYQPLEGSLMINK 246
Db 98 -----NCPRESSASFLSRIFMWTGLIVRGYQPLEGSLMINK 138
Qy 247 EDITSEQVVPVLVKNMKKECAKTRKQPKVYVSSKDPQPKSSKYDANEVEALIVKSPQ 306
Db 139 EDITSEQVVPVLVKNMKKECAKTRKQPKVYVSSKDPQPKSSKYDANEVEALIVKSPQ 198
Qy 307 KENNPSLFKVLKTFGEYFLMSFFPKAIHDIIMFSGQILKLIKPVNDYKAPDMQGYFY 366
Db 199 KENNPSLFKVLKTFGEYFLMSFFPKAIHDIIMFSGQILKLIKPVNDYKAPDMQGYFY 258
Qy 367 TVLLFVTACIQTVLHQYFHI CFVSGMR IKTAVIGAYYRKALVITNSARKSSTYGEIYNL 426
Db 259 TVLLFVTACIQTVLHQYFHI CFVSGMR IKTAVIGAYYRKALVITNSARKSSTYGEIYNL 318
Qy 427 MSVDAQRFMDLATYINNIWSAPLQVILALYLLMNLGPSVLAGAVANVLPVNAVMAWK 486
Db 319 MSVDAQRFMDLATYINNIWSAPLQVILALYLLMNLGPSVLAGAVANVLPVNAVMAWK 378
Qy 487 TKTYOVAMKSKDNRIKLMEIILNGIKVLKYAMELAFKQVLAIRQBELVLRKSAVLS 546
Db 379 TKTYOVAMKSKDNRIKLMEIILNGIKVLKYAMELAFKQVLAIRQBELVLRKSAVLS 438
Qy 547 AVGFITWCTPFLVALCTFAVYVITDENNILDQTAVYSALFPIILFPLNILEPVYISI 606
Db 439 AVGFITWCTPFLVALCTFAVYVITDENNILDQTAVYSALFPIILFPLNILEPVYISI 498
Qy 607 VQASVSLKRLRIFLSHELEPDSIERRPVODGGTNSITVRNATFTARSDDPLNTGTF 666
Db 499 VQASVSLKRLRIFLSHELEPDSIERRPVODGGTNSITVRNATFTARSDDPLNTGTF 558
Qy 667 SIPEGALVAVVGQVCGKSSLSALLAEMKVEGHVAKGSVAAYVPOAWI QNDLRENI 726
Db 559 SIPEGALVAVVGQVCGKSSLSALLAEMKVEGHVAKGSVAAYVPOAWI QNDLRENI 618
Qy 727 LFGQLEBPYRYSTIQACALLPDLLEIIPSGDRETEIGKGVNLSGGQKQRYSLAAVYSNA 786
Db 619 LFGQLEBPYRYSTIQACALLPDLLEIIPSGDRETEIGKGVNLSGGQKQRYSLAAVYSNA 678
Qy 787 DIYLFDDPLSAVDPAHVGHIFENVIGPKMKKTRILVTHSMGYLPQVDV IYMSGGKI 846
Db 679 DIYLFDDPLSAVDPAHVGHIFENVIGPKMKKTRILVTHSMGYLPQVDV IYMSGGKI 738
Qy 847 SEMGSYQELLARDGAFELRTYASTEQEODAEENGVTGVS GPKKAKOMENGLVYDSA 906
Db 739 SEMGSYQELLARDGAFELRTYASTEQEODAEENGVTGVS GPKKAKOMENGLVYDSA 798
Qy 907 GKQLORLSSSSSYSGDISRHHNSTAELOKAEKKEETMKLEADKQGVKLSVYDYM 966
Db 799 GKQLORLSSSSSYSGDISRHHNSTAELOKAEKKEETMKLEADKQGVKLSVYDYM 858
Qy 967 MKAIGLISFLSLIFLFCNHYASALASNYWLSLMTDDPIVNGOEHTKYRLSVYGA LGISQ 1026
Db 859 MKAIGLISFLSLIFLFCNHYASALASNYWLSLMTDDPIVNGOEHTKYRLSVYGA LGISQ 918
Qy 1027 GIAVFGYMAVSIIGIILASRC LHVLDLHLSIRSPMSPEPTPSGNLVNRFSEKLDVTYSM 1086
Db 919 GIAVFGYMAVSIIGIILASRC LHVLDLHLSIRSPMSPEPTPSGNLVNRFSEKLDVTYSM 978
Qy 1087 IPEVIKFMGSLFNVIGACIYI LLA TP IAA I I I P P L G L I Y F V O R F Y A S S R O L K R L E S V 1146
Db 979 IPEVIKFMGSLFNVIGACIYI LLA TP IAA I I I P P L G L I Y F V O R F Y A S S R O L K R L E S V 1038
Qy 1147 SRSPPVSHFNETLIGVSVIRAFEEOEERFIHOSDLKVDENOKAYYPSIVANRWLA VRLECV 1206

Db 1039 SRSPVSHFNHETLLIGSVITAFEEQERFIHOSDLKVDENQKAYPSIVANFMLVRLCEV 1098
Qy 1207 GNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIIVAVERLKE 1266
Db 1099 GNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIIVAVERLKE 1158
Qy 1267 YSETEKAPMOIQETAPPSWPQVGRVVEFNRYCLRYREDDLPVLRHINVTNGEKGIV 1326
Db 1159 YSETEKAPMOIQETAPPSWPQVGRVVEFNRYCLRYREDDLPVLRHINVTNGEKGIV 1218
Qy 1327 GRTAGAGSLLTGLFRINESAGEIITDGINIAKIGHDLAKFKITIIPODVLSSGSLRM 1386
Db 1219 GRTAGAGSLLTGLFRINESAGEIITDGINIAKIGHDLAKFKITIIPODVLSSGSLRM 1278
Qy 1387 NLDPFSQYDEBEWVTSLELAHKDFVSALPKLDHECAGGENLSVGOQLVCLARALLR 1446
Db 1279 NLDPFSQYDEBEWVTSLELAHKDFVSALPKLDHECAGGENLSVGOQLVCLARALLR 1338
Qy 1447 KTKILVIDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTR 1496
Db 1339 KTKILVIDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTR 1388

RESULT 14
US-09-939-853A-88
; Sequence 88, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OR INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 1528
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-88

Query Match 71.9%; Score 7002.5; DB 3; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

Qy 1 MALRFGSADGSDPLMDMNVNTWNTSNPDFTKCFQNTVLVWPCFYLMACFPEYFLYLSRH 60
Db 1 MALRFGSADGSDPLMDMNVNTWNTSNPDFTKCFQNTVLVWPCFYLMACFPEYFLYLSRH 60
Qy 61 DRGYOMPLUNKTALGFELMIVCMADLFYSFWRSGQVLAPLALVSPITLGIITLLA 120
Db 61 DRGYOMPLUNKTALGFELMIVCMADLFYSFWRSGQVLAPLALVSPITLGIITLLA 120
Qy 121 TFLIOLERRKGVQSSGIMLTFMLVALVQALAIRSKITMTALKEDAQVDFRDIITYVYFS 180
Db 121 TFLIOLERRKGVQSSGIMLTFMLVALVQALAIRSKITMTALKEDAQVDFRDIITYVYFS 180
Qy 181 LLLIQLVLSGFSRDSPLFSETIHDNPPCESSASFLSRITFWMITGLIVRGYRQPLEGSD 240
Db 181 LLLIQLVLSGFSRDSPLFSETIHDNPPCESSASFLSRITFWMITGLIVRGYRQPLEGSD 240
Qy 241 LMSLINKETSRQVAVVLYKMKKECAKTRKQPVKYVYS-SKDPAPRKSSKTDAAEEVEA 299
Db 241 LMSLINKETSRQVAVVLYKMKKECAKTRKQPVKYVYS-SKDPAPRKSSKTDAAEEVEA 299

Qy 300 LIYSPQKEMNPSEFKVLYTFCGPFLMSFFFKALHDLMPFSGRQILKLKIFVNDTAP 359
Db 301 LIYSPQKEMNPSEFKVLYTFCGPFLMSFFFKALHDLMPFSGRQILKLKIFVNDTAP 360
Qy 360 DMQGFYTVLLFVYACIOTLVHQQYFHCPSGRIKTAIVGAYRKALVITNSARKST 419
Db 361 DMQGFYTVLLFVYACIOTLVHQQYFHCPSGRIKTAIVGAYRKALVITNSARKST 420
Qy 420 VGEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLLMLNLPVSLAGVAVVLMVPV 479
Db 421 VGEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLLMLNLPVSLAGVAVVLMVPV 480
Qy 480 NAYNAKTKTYQVAHMSKONRILKANEIINGIVLYKLYAMELAFKQVLAIRQELKVL 539
Db 481 NAYNAKTKTYQVAHMSKONRILKANEIINGIVLYKLYAMELAFKQVLAIRQELKVL 540
Qy 540 KKSAYLAVGTFVWCPELVALCTPAVYVYIDENNILDAOTAFVSLAFNILEPPLNLT 599
Db 541 KKSAYLAVGTFVWCPELVALCTPAVYVYIDENNILDAOTAFVSLAFNILEPPLNLT 600
Qy 600 PMVSSIVQASVSLKRLRIFLSHELEBDSIERSPVXDGGTNSITVNAATFTWASDDP 659
Db 601 PMVSSIVQASVSLKRLRIFLSHELEBDSIERSPVXDGGTNSITVNAATFTWASDDP 659
Qy 660 TLNGITTSIPRGALVAVYVGYGCGKSSLSALLAMBQKVEGHAIKSVAVYVPOAMIQN 719
Db 660 TLNGITTSIPRGALVAVYVGYGCGKSSLSALLAMBQKVEGHAIKSVAVYVPOAMIQN 719
Qy 720 DLSRENILFGQULEEPPYRSVIOACALLPDEIILPSGDRTEIGEGVNLSSGQKORVSLA 779
Db 720 DLSRENILFGQULEEPPYRSVIOACALLPDEIILPSGDRTEIGEGVNLSSGQKORVSLA 779
Qy 780 RAYVSNADIVYFDPLSAVDAVHGKHIFENVIGRGMKKNKTRILVTHSMYSLPQVDYI 839
Db 780 RAYVSNADIVYFDPLSAVDAVHGKHIFENVIGRGMKKNKTRILVTHSMYSLPQVDYI 839
Qy 840 VMSGKISEMWSYOELLARDGAPLEPLTYASTQEBDAEENGVTGVSQPKKAKOMENG 899
Db 840 VMSGKISEMWSYOELLARDGAPLEPLTYASTQEBDAEENGVTGVSQPKKAKOMENG 896
Qy 900 MLNVDASQKOLOROLSSSSSGDISRHHNSTAELOKAEKFTWKLMEADKQGTGVK 959
Db 900 MLNVDASQKOLOROLSSSSSGDISRHHNSTAELOKAEKFTWKLMEADKQGTGVK 955
Qy 997 MLVDTVQKHLQRLHNSSSHSGTSSQSHSIAELQKGA-KEETWKLMEADKQGTGVQ 955
Db 997 MLVDTVQKHLQRLHNSSSHSGTSSQSHSIAELQKGA-KEETWKLMEADKQGTGVQ 955
Qy 960 LSVYMDYKAIGLTISFLSIFLPMCNHVSALASNYMLSLMTDD-PIYNGTOEHTKVLRSV 1018
Db 960 LSVYMDYKAIGLTISFLSIFLPMCNHVSALASNYMLSLMTDD-PIYNGTOEHTKVLRSV 1015
Qy 956 LSVYMDYKAIGLTISFLSIFLPMCNHVSALASNYMLSLMTDD-PIYNGTOEHTKVLRSV 1015
Db 956 LSVYMDYKAIGLTISFLSIFLPMCNHVSALASNYMLSLMTDD-PIYNGTOEHTKVLRSV 1015
Qy 1019 YGALGISGIAVFGYSMAVSTGGILASRCULHVDLHSILRSPMSFFERTPGNLYNRFSK 1078
Db 1016 YGALGISGIAVFGYSMAVSTGGILASRCULHVDLHSILRSPMSFFERTPGNLYNRFSK 1075
Qy 1079 ELDTVDSMIPRIVIKFMSGLFNVIGACVILLATPIAIIIPGLIYFPQRPVVASR 1138
Db 1076 ELDTVDSMIPRIVIKFMSGLFNVIGACVILLATPIAIIIPGLIYFPQRPVVASR 1135
Qy 1139 QKRLLESRSRPVYSHFNHETLLIGSVITAFEEQERFIHOSDLKVDENQKAYPSIVANRW 1198
Db 1136 QKRLLESRSRPVYSHFNHETLLIGSVITAFEEQERFIHOSDLKVDENQKAYPSIVANRW 1195
Qy 1199 LAVRLCEVGNCTVLPALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNI 1258
Db 1196 LAVRLCEVGNCTVLPALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNI 1255
Qy 1259 VAVERLKEYSYTEKAPMOIQETAPPSWPQVGRVVEFNRYCLRYREDDLPVLRHINVTIN 1318
Db 1256 VAVERLKEYSYTEKAPMOIQETAPPSWPQVGRVVEFNRYCLRYREDDLPVLRHINVTIN 1315
Qy 1339 GGEKVGIVGRTGAGKSSLLTGLFRINESAGEIITDGINIAKIGHDLAKFKITIIPODPV 1378
Db 1316 GGEKVGIVGRTGAGKSSLLTGLFRINESAGEIITDGINIAKIGHDLAKFKITIIPODPV 1375
Qy 1379 LFSGSLRMLNDPFSQYDEBEWVTSLELAHLKDFVSALPKLDHECAGGENLSVGOQLV 1438

Db 1376 LPSGLRNKLDPSQSDSEEWMALELAHKGFSVALPDKLNHECAEGENLSVGQRQLV 1435
QY 1439 CLAAALLRKTKILVLDENTAVIDETDILLIOSTIRTOEDCTVLTIAHRLNTIMDYTRVI 1498
Db 1436 CLAAALLRKTKILVLDENTAVIDETDILLIOSTIRTOEDCTVLTIAHRLNTIMDYTRVI 1495
QY 1499 VLDRGEIOEGAPSDILQORGLFYSMAKAGLV 1531
Db 1496 VLDRGEVRECGAPSELLQORGLFYSMAKAGLV 1528

RESULT 15
US-10-618-281-63
; Sequence 63, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Esteell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously Known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 1303
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-618-281-63

Query March 62.3% Score 6068.5; DB 5; Length 1303;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1229; Conservative 17; Mismatches 57; Indels 207; Gaps 10;

QY 22 WNTSNPPTKCFQVTVLWVPCFYLMACFPYFLYLSHRDGYIOMTPLNKTKTALGFL 81
Db 1 WNTSNPPTKCFQVTVLWVPCFYLMACFPYFLYLSHRDGYIOMTPLNKTKTALGFL 60
QY 82 WIVCMADLFYSFWRERSGIFLAPVPLVSPILLGITTLATLTLQLERRKGVQSSGIMLTF 141
Db 61 WIVCMADLFYSFWRERSGIFLAPVPLVSPILLGITTLATLTLQLERRKGVQSSGIMLTF 120
QY 142 WLVALVCAALALRSKIMTALKEAQNVDLFRDITFYVVSILLIQLVLSCEFSDRSPLSET 201
Db 121 WLVALVCAALALRSKIMTALKEAQNVDLFRDITFYVVSILLIQLVLSCEFSDRSPLSET 177
QY 202 IHDNPPRESSASLSRITTFWMTGLVYRGROPLEGSDLSLNKEDPSEGVNPLVKNW 261
Db 178 IHDNPPRESSASLSRITTFWMTGLVYRGROPLEGSDLSLNKEDPSEGVNPLVKNW 237
QY 262 KKECAKTRKQPVKVVSSKODAPKESKVDANEVEALLVKSPOKEMNPSLFRVLYKTF 321
Db 238 KKECAKTRN-----SSGSGESCSANTALF-----PA--PTCHKSF 271
QY 322 GPFLYMSFFPKAIDHLMFSGPOLIKLIRFVNTKAPDMOGYFYTYLLFTTACLTQTLVL 381
Db 272 QALSLL-----LCRLIKFVNDTKAPDMOGYFYTYLLFTTACLTQTLVL 314
QY 382 HOYHICFVSGMRKITAIVIGAVYRKALVITNSAKSSSTVGEIVNIMSGVDAORFMDLATYI 441
Db 315 HOYHICFVSGMRKITAIVIGAVYRKALVITNSAKSSSTVGEIVNIMSGVDAORFMDLATYI 374
QY 442 NMITSAPLOVTLAIVLMLNIGSVLAGVAVMLVMPVNAVAMKTKTYOYAHKSKDNR 501
Db 375 NMITSAPLOVTLAIVLMLNIGSVLAGVAVMLVMPVNAVAMKTKTYOYAHKSKDNR 416
QY 502 IKLMNELLNGIKVLKVLAMELAFDKVLAIRQBELKVLKKSAYLSAVGTFTWCTPFLVA 561

Db 417 IKLMNELLNGIKVLKVLAMELAFDKVLAIRQBELKVLKKSAYLSAVGTFTWCTPFLVA 476
QY 562 LCTFAVYVTTDENNTLDAQTAFFVSLAFNIRPFLNTPMWTSSIVQASVSKRLRIFLS 621
Db 477 LCTFAVYVTTDENNTLDAQTAFFVSLAFNIRPFLNTPMWTSSIVQASVSKRLRIFLS 528
QY 622 HEELEPISIBRRPKDGGGNTSIVRNATPFWASDPTLNGIFESIPEGALVAVGVG 681
Db 529 GATSERGPWGRPKHG-----TRQASFSVAPGVLCRPSITFSIPEGALVAVGVG 581
QY 682 CGKSLLSALLAEWDKVEGHVAIKGSVAYVYPOQAMIQNDLSRENILFGCQLEBPYRSVI 741
Db 582 CGKSLLSALLAEWDKVEGHVAIKGSVAYVYPOQAMIQNDLSRENILFGCQLEBPYRSVI 641
QY 742 QACALLPDLBTLPSSGDRTEIGEKVNLSSGQKQVSLARAVYSNADITYLPDPLSAVDAA 801
Db 642 QACALLPDLBTLPSSGDRTEIGEKVNLSSGQKQVSLARAVYSNADITYLPDPLSAVDAA 701
QY 802 VGKHFENNVIGPKMLNKTRILVTHSMSTLPQVDVLIWVSGKISEMGSYQELLARDGA 861
Db 702 VGKHFENNVIGPKMLNKTRILVTHSMSTLPQVDVLIWVSGKISEMGSYQELLARDGA 724
QY 862 FAELFRYASTEQDQDAEENGVTGVSQPKAKQKQNGMLVTDSAGKQLQRLSSSSYS 921
Db 725 -----S 725
QY 922 GDISRRHNSTAELQKAEKKEETKMEADKQOYKLSYVMDYKALIGLIFSLIFL 981
Db 726 CDL-----QYKLSYVMDYKALIGLIFSLIFL 753
QY 982 FMCNHSALASNYTLSTMTDPIYNGTOETKRLSYGALGISQGIYAVGYSNAVSIGG 1041
Db 754 FMCNHSALASNYTLSTMTDPIYNGTOETKRLSYGALGISQGIYAVGYSNAVSIGG 813
QY 1042 ILASRCHVDLHLSILSPSPFERTPSGNLVNRSKELPTVDSMTEPVIKMFPSLFPV 1101
Db 814 ILASRCHVDLHLSILSPSPFERTPSGNLVNRSKELPTVDSMTEPVIKMFPSLFPV 873
QY 1102 IGACIVILLATPIAIIIPPLGLIYFPVQFYVASSRQLRLSEVSRSPPVSHNETLLG 1161
Db 874 IGACIVILLATPIAIIIPPLGLIYFPVQFYVASSRQLRLSEVSRSPPVSHNETLLG 933
QY 1162 VSVTRAEBOERPHQSDLVKVDENQKAYYSIVANRWLAIRLBCVNCIYLPALPAVIS 1221
Db 934 VSVTRAEBOERPHQSDLVKVDENQKAYYSIVANRWLAIRLBCVNCIYLPALPAVIS 993
QY 1222 RHLISAGLVGLSVSYLOVTVLNMVLRMSSEMETNVAVERLKEVSETEKEAPWJOET 1281
Db 994 RHLISAGLVGLSVSYLOVTVLNMVLRMSSEMETNVAVERLKEVSETEKEAPWJOET 1053
QY 1282 APPSPWPQVRVERPNCLARYEDLPVLNHNVTINGEKVGIYGRTGAKSSLTJGLF 1341
Db 1054 APPSPWPQVRVERPNCLARYEDLPVLNHNVTINGEKVGIYGRTGAKSSLTJGLF 1113
QY 1342 RINESAGEIILIDGINIAKIGLHDLRFXITIIPODPVLPFGSLRMLNIDPFSQYSDSEVMT 1401
Db 1114 RINESAGEIILIDGINIAKIGLHDLRFXITIIPODPVLPFGSLRMLNIDPFSQYSDSEVMT 1173
QY 1402 SLELAHLKDFVSALPDLHDECAEGENLSVGQRQVLCARALLRKTKIIVLDEATYAVD 1461
Db 1174 SLELAHLKDFVSALPDLHDECAEGENLSVGQRQVLCARALLRKTKIIVLDEATYAVD 1233
QY 1462 LETDDLIQSTIRTOEDCTVLTIAHRLNTIMDYTRVIVLDRGEIOEGAPSDILQORGLF 1521
Db 1234 LETDDLIQSTIRTOEDCTVLTIAHRLNTIMDYTRVIVLDRGEIOEGAPSDILQORGLF 1293
QY 1522 YSMAKAGLV 1531
Db 1294 YSMAKAGLV 1303

Mon Dec 19 12:36:09 2005

us-10-665-283-4.rapbm

Page 16

Job time : 165.846 secs

Oy	495	MKSXONRITKLMNELLINGKVKULVAMELAFQDKYALAEQOEELKVKYKCAVYSAVTFPMV	554
Db	1	MROGDSRARLTSILNRSKTIKHFHGBEAFIDRVUGIRGQELGALRTSGLFSVSYSFQ	60
Oy	555	CTPPLVALCTFAVYVTTIDENNILDAQTAQVSLALFNILRPPLNIPMWISSIVQASVSLK	614
Db	61	VSTFLVALVFAVHTLVAE-NANNAEKAFVTLVTYNINKAAQAFPEFISHIVQARVSFD	119
Oy	615	RLRFLSHEELPSPSIERRPVKOGGNTSIVRNATPFMASSDPTLNGITPFIPEGALV	674
Db	120	RLVYFLCLEBVDPEGVDBSSSGSAAGKDCITIGATPFMSGSESPCLHINULVYQGLL	179
Oy	675	AVVQGVGGCKSSLISALLAEMDKYEGHVALKGSVAAYVPQAMIONDSLRENILPFCQLEB	734
Db	180	AVVPVPVAGKSSLISALLGELSKYEGFYSIGEAAYVPOEMAMONTSVENAVCFCQELDP	239
Oy	735	PYRYSVIOACALPLELPEGSDTEIGEKEVNLSSGGCKQVSLAPARAYSNADYILFPDP	794
Db	240	PMLERVLEACALQGDVDSFPEGHITSIGEQGMNLSSGGCKQSLSLARAYRRRAAYILDPD	299
Oy	795	LSAVDAHVGHKIFENVTIGPKOMLKNKTRILVTHSMSTYLPQVDVILVNSGGKISEMGSYOE	854
Db	300	LAALDAHVGHGVNQVIGPGLLOGTTRILVTHALHILPQADWILVILVLANGIAENGSYOE	359

QY 855 LLARDGAFAEFLRYASTEOBDAENGVTGSGPKEAKOMENGLVTSAGK--QLOK 912
DB 360 LLQRKGLVCL-----DQARQGDG--EGETEGTSTKDPG-----TSAGRPBLRL 407
QY 913 QLSSSSSYSGDISHNHSTAELOCAEAKET---WKLMEADKAQCGQVLTSTYMDYK 968
DB 408 ERSKTS-----VPEKDRITSQAQ--TEVPLDPPDRAGPAGK--DSIOGRVATVHLAYLR 460
QY 969 AIGFISFISFIFEMCNHVSALASNYMLSLMTDDPIYNGTOEHKTRLSVYGALISQGI 1028
DB 461 AVGPPLCYALFELCQOVASFCCGYMLSLMADBPANGGQTOAALRGITGLIGCLOAI 520
QY 1029 AVFGYSAVNSIGGILASRCLHVDLHSLTRSMPSEFFETPSGNTVNRPSKELDVDSMP 1088
DB 521 GLFASMAVALLGARGASRLPQRLMDVRSPISEFFETPIGHLNRSKEDTDVVDIP 580
QY 1089 EVIEMFMSLEPNVIGACIVILLATPIAIIIPRGLIYFPQRFVASSRQIKLESISR 1148
DB 581 DKRLSLMYAFGLLEVSILVAVATPLATVAILPLFLYAGPQSLVYVSSCOLRRLESASY 640
QY 1149 SPVYSHNETLLGVSVIRAFEEORFIIHQSLDKDENOKAYYPSIVANRMLAVRLCYGN 1208
DB 641 SSVCSHMAETQGSTVVRARFTQAPFVAQNNARVDESQRISEFPLVADRMILAAVLLGN 700
QY 1209 CIVLFAALFAVISHRSLSAGLVGHSVSYSLQVTTYLMLVMSSEMETNIVAVERLKEYS 1268
DB 701 GLVFAALTCAVLSKHLHSAGLVGSVSAALQVOTQLQVAVNRNMVDLENSIYSVERMODYA 760
QY 1269 ETEKEAPMOIOETAPPSWPOVGRVERPNYCLRTREDDFVLRIHINTINGEKYGIYGR 1328
DB 761 WTPKEAPRRLPTCAQAPWPQGGQIEFRDFGLRYAPBELPLAVQGVSFTHGEKGIYGR 820
QY 1329 TGAKSSSLTGLFPIINESAGEIITIDGINIAKIGHDRPKITTIPODPVLFSGSLRML 1388
DB 821 TGAGKSSLASGLIKIOEABEGIMIDGVPDAHVGHTLRSKISIIPODPILFPGSLRML 880
QY 1389 DPFQYSDDEEWTSLLEHLAKDPFVSALPDKLHCEACBEGENLSVGORQVLCARALAKRT 1448
DB 881 DLDEHSEALWALLETYQLKALVASLPGQLOYKCADGEBDSVQKDLCLARALAKRT 940
QY 1449 KILVDEATAVDELTDLIQSTTRTOEDCTVLTIANRLTINDTYVILDKGEIOY 1508
DB 941 QILILDEATAVDELTDLIQSTTRTOEDCTVLTIANRLTINDTYVILDKGEIOY 1000
QY 1509 GAPSLLQORGLFYSMADAGLV 1531
DB 1001 GSPAQLLAQKGLFYRLAQESGLV 1023

RESULT 2
US-11-090-439-24
Sequence 24, Application US/11090439
Publication No. US20050266442A1
GENERAL INFORMATION:
APPLICANT: Squillace, Rachel
APPLICANT: Weiner, Michael P.
TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
FILE REFERENCE: 24318-502
CURRENT APPLICATION NUMBER: US/11/090,439
PRIOR FILING DATE: 2005-03-25
PRIOR APPLICATION NUMBER: 60/556,344
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 1581
TYPE: PRF
ORGANISM: Homo sapiens
US-11-090-439-24

Query Match 21.7%; Score 2115; DB 7; Length 1581;
Best Local Similarity 31.9%; Pred. No. 66-147;

Matches 527; Conservative 305; Mismatches 564; Indels 256; Gaps 37;
QY 32 CFQNTVLVWVPCFYLMACPPFFYL--YLSRHDGTYQMTPLNKTALGF-----LMTWICW 86
DB 26 CFVDALNVVAVHVLFTTFFILPILFGMSQSSKVIHHS-----TWLHFPCHNLRLWLTFF 79
QY 87 ADLEFSEWERSRGIFLAPV-----FLVSPT---LIGITTLATPFILOERRKVOSSG 136
DB 80 MLFVLVCEIAEGLSDGVESHHLHLYMPAGMAFMAVTSV--YYHNIEITSNFKLLI 137
QY 137 IMTLFVLVLCALAILRSKIMTKEDAQVDLFRDTF-----YVPELLILQVLVS 189
DB 138 ALVYVWTLAFL-----TTIKFVKFLDHAIGSQRFCLTGLLVILYGLMLIVE--VN 188
QY 190 CFSDRSLPSETHIDPNCPESS-----SASFSLRTTFMWTGLYRGYROPLEG 238
DB 189 VIRRRRIYFFKTRPEVPR--BEDDODLGRPLQPPVNLISKOTYMMMAFIYAKKPI-- 245
QY 239 SDLMSLNKEDTSQGVNVLVYK--NMKECAKTRKQPVKVVYSKDPAPQKRESSKVDANE 295
DB 246 -DLRAIGK-----LPLMARLNTNYQRLC-----E 268
QY 296 EVELLIYKSPQ--KEMNPSLFVULYKTPGYFLMSFFPKAIDHLMFSGPQILKILIFV- 353
DB 269 AFDAQVRKDIQGTQABAIWQALSHAFGRRLVLSSTERIIDLGLFAGPLCIPEIYVHLG 328
QY 354 --NDTKAPDMQ-----GYFYVLLFVTRACLOTVLVHQVFHICFVSGMRK 396
DB 329 KENVVPOKTOFLGVYVSSQEFANNAVLAFLVLLIORTLOASYYAIAETGINLR 388
QY 397 TAVIGAVYRKA--LVITNSARKSSVGEIYNLMSVDAORFMDLATYINMTWSAPLOVILA 454
DB 389 GAIQTKIYNKIMHLSNTLSNGEATAGQICNLVAIDTNQMLWFFPLCPNLWAMPVQIIVG 448
QY 455 LYILMLNLGSPVLGAVVMTLVVNVAVMMAMKTTQYVAHKSQDNRIKLMNBLINGKV 514
DB 449 VILLYYLIGVSALGAAVYIILAPVOYFVAFKLSQAQRSTLEYENERIKQTNEMLRGKL 508
QY 515 LKLYAMELAFQDKVLAIRQEBELKVLKKSAYLSAVGFTWVCTPLVALCTPAVYVTTIDEN 574
DB 509 LKLYAMENIIFTRVETTRKRMKMTSLRAPAIYTSISIFMNTAIPILAUVITVGHVSFFKE 568
QY 575 NILDAQAFVSLALFNILRPPLNLIIPWISSIVQASVSLKRLIFLSHEELPEPSI---- 630
DB 569 ADFSPVAFAFLSLFHLVLPFLSSVSSTVAVLVSVQLSFSLSAEIRBQCAPHB 628
QY 631 -----ERRPVK-----DGGTNY-SITYVNAATYTW 653
DB 629 PTPQGPASKYQAVPLRVNRRKRPAREDCRGLTGPLQSLVPSADADADNCVQINGGYFTW 688
QY 654 ARSDPPTLNGITFSIPREGALVAVVGVGCGKSSLSLALLAEMDKVEG-----700
DB 689 TPDGIPILSNITITIPRQQLTMIVGVGCGSSLSLALLAGEMOKYSGAVFWSSLPDSIEIG 748
QY 701 -----HVAIKGSVAAYVPOQAMIONDSIRENITLFGCOLLEBYYSVIOACALL 747
DB 749 EDPSPERATADLDIRKGPVAVASQKRWMLNATVEENIIESPFNKORYKVMVIEACSLQ 808
QY 748 PDLBILPESGDRTEIGEKVNSGGQKQVSLARAVYSNADLYLPDDPLSAVDAYKGIIF 807
DB 809 PDIDILPHGDQTOIGERGINSGGQRORISVARALYOHANVAFDDPFSALDIHLSDIIM 868
QY 808 ENVIQPGKMLK-----NKTIRLVTHMSNYLQOVVYIYMSGSKIEMSGSYOELLARDGAF 862
DB 869 Q-----AGITLRLDDKKTIVLVTHKQLYLPHADMTIMKQGTIQRBESTLQFORSEOL 923
QY 863 AEFLRTYASTEOBDAENGVTGVSFGPEKAKQOMENGLVTSAGKQLOROLSSSSSYSG 922
DB 924 FEHMKITLMN--ROQDELEKEIVT-----ERKATPEPQG-----LSRAMSSSDGL-- 965
QY 923 DISRHHNSTAELOCAEAKKETWKLMEADKA-----QTGOYKLSYVWDYKALGLFISFL 977
DB 966 -----LQDEEEEEEBAESEBDDNLSMLHQBAREIPWRACAKYLTSSAGITLNL 1014

```

QY 978 SIFLEMCNCHVASLASNMYLMSLTDDPIV-----NCGOEH-----KYRLSYGALG 1023
Db 1015 LVFSOLKHMVLVALDIPMLAKTBDALTITTPAARCISJQECTLDQTYAMFTVLCSLG 1074
QY 1024 ISQGIATVGYMAVASIGGILASRCLHVDLHSLIRSPMSFFERTPSGNLVNRFSEKLDTV 1083
Db 1075 I---VLICVTSVTEWMTGKVKAKRLHRSILNRNII LAPRPFETPTLGSILNRFSSDCNTI 1131
QY 1084 DSMIEPVILKMFPGSLFNVIACIVLLATPLIAITIIPLGLIYFFVQRFYVAASQOLKRL 1143
Db 1132 DQHPSTLECSRSTLLCVSALAVISVYTPVELVALLPLAIVCYPIQYCFRFRASHDLOOL 1191
QY 1144 ESVSRSPYASHNETILLGVSIVLRAFEEOERFHHOSDKVDENOKAYPSIVANRMLAVRL 1203
Db 1192 DDTTOLPLSHFAETVEGTLTTRAFRYEARFOQKLETTDSNNIASLFLTAANRMLVYRM 1251
QY 1204 ECVNGCIVLFAALFAVIS--RHSLSAGLVGSVSYSLQVTTYLMLVMSSEMETNIYAV 1261
Db 1252 EYIGACVVLIAVTSISNSLHRELSAGLVGGLTVALVWVSNYLNMMVRNLDLMELQGLAV 1311
QY 1262 ERLKYESETEK-----APMOIETAPSPSPVQGYVEFPRNYCLRYEDEDVFLRHIV 1315
Db 1312 KRIMHMLTKEASESYEGLAPSLI-----PKMPPDQKXIOQNLVRYSDSSLKPYLKHVNA 1366
QY 1316 TINGEKGIVGARTGAGKSSLLTGLFRINESAGEIIIDGINIAKIGHLDLRFKTLTIPO 1375
Db 1367 LISPOKIGICGRTSSGKSSSFLAFRRVNDTFEGHIIIDGIDIALPLHTIRSRSLILO 1426
QY 1376 DPEVFSGLRMLNLDPEFSQSDSEWVTSLELAHKDQFVSAIPDKLDHCEAGGEMNLVSGOR 1435
Db 1427 DPEVFSGTIRFNLDPERKCSDBTLEWALEIQLKLVKALPGGDAIITEGGENPSSQGR 1486
QY 1436 QLVCLARALKRKTILVDERTAAVDLETDDLIGSTITFOBEDCVLTITLARLTINDYT 1495
Db 1487 QLFCLARAFVTRKTSFFIMDEATASIDAMATENILOKVWMTAADRIVTITARVHTILSAD 1546
QY 1496 RVIYLDKGEIOEYGAPSDILOOR--GLFSPMAK 1526
Db 1547 LVYIVAKRAIIIEFDKPEKILSRKDSVFASFVR 1578

RESULT 3
US-10-995-561-527
; Sequence 527, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 527
; LENGTH: 1588
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-527

```



```
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 969
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-969

Query Match          10.0%; Score 969; DB 6; Length 410;
Best Local Similarity 50.8%; Pred. No. 9,8e-64;
Matches 197; Conservative 66; Mismatches 121; Indels 4; Gaps 2;

Qy 495 MSKDNRIKLMNEILNGIKVLYAMELAFKDKVLAIRQEBLVKKSAYLSAVGTFTWV 554
Db 1 MRQDSRAKLTSSILRNKSKITKFGHEGAFIDRVLGIGQELGALRTSGILFVSLSVFO 60
Qy 555 CTPELVALCTPAVYVTTIDENNILDAQTAFVSLALFNILRFPNLIPWVSSIVQASVSLK 614
Db 61 VSTFLVALVFAVHTLVAE-NAMNAEKAFTVLTVINILINKAQAFLPFIHSIVQARVDFD 119
Qy 615 RLRLFLSHLEPEPDSIERPKVYDGGGTSITYRNAFTFMASDPPTLNGITFSIEGALV 674
Db 120 RLVLFLCLIEEYDPGVADSSSGSAGKDCITIQSTFAMSOQSPCLHRINLTVQGGCL 179
Qy 675 AVVGQVGKSSLSLALAEADKVEGHAIKGSVAVVPOQAWIONDSIRENLIFFCOLEE 734
Db 180 AVVGQVGKSSLSLALAEKLSKVEGVSISGAVALVPOEAVVQTTSVENECFQOEIDP 239
Qy 735 PYRSVIOACALLPDLLEILPSSDRTEIGKGVNLGGQKORVSLARAVYSNADIVLPDP 794
Db 240 PMLERVLEACALQPVVDSPBSIHHSIGQGWNLSGGQKORSLARAVYRKAAVLLDDP 299
Qy 795 LSAVDAYGKHIFENVIGPKGMLKNKTRILVTHSNYSYLPQVDVVIYMGSGKISEKGSYOE 854
Db 300 LAALDAHYGHVFNQVIGPGGILQCTRIILVTHALHILPOADWIIIVLANGALAEKGSYOE 359
Qy 855 LLARDGAFAEFRTASTTEQBDARENG 882
Db 360 LLQKRGALVCLD--DQARQPEDRGSGG 384

RESULT 7
US-10-613-744-14
; Sequence 14, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 14
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-613-744-14

Query Match          9.1%; Score 883.5; DB 6; Length 407;
Best Local Similarity 48.7%; Pred. No. 1.8e-57;
Matches 167; Conservative 74; Mismatches 95; Indels 7; Gaps 4;

Qy 1551 LTRLAEPAPBRY-----RTERRARFVSKGNCVAHKNIHQEQPLQDVPTLVLDKWP 1606
Db 16 LPKQARDLQ-PRHISRDPTKAKIQIVRKQDKCNHGNVETRYRLTDTITLVLDKMR 74
Qy 1607 HTLLIFTMSPFLCSWILFAMVWMLIAFAHGLDPAPEGTN-VPQVTSIHSFSSAPFSLIEVQ 1665
Db 75 FNLIIFVWVTVVWLPFGMIWMLIAYIRGMDHIEDSPWPCVTNNGFVSAPFSLIETE 134
Qy 1666 VTIGEGGAMTEECPLAILIIVQNIYGLMNAIMLGCIFFKTNQAHRAETLLPSKNAV 1725
Db 135 TTIGGYRVITDKCPREGIILLIOSVLSGSIYNAFMVGCMPFKIIOPKKRAETLVFSTAV 194
Qy 1726 ITLRHGRLCFMLRVGDLRKSMIISATIMQVVRKTSPEGEVPLHQVDIMENGVGANG 1785
Db 195 ISMRDGLKCLMFRVGDRLNSHIVASIRAKLIKQTSBGEFPLNQTDINVGYYTGDR 254
Qy 1786 IFVLAPLIYHVIDNSPLVDLAPSDLHHODEIIVLBSGVETGTTTQARTSYLAD 1845
Db 255 LFIVSPILSHIEINQGSFWEISKAQL-PKELEIIVLIEGVATELQWTCQARSYINSE 313
Qy 1846 ILMGQRPVPIAEDGRYSVDYSKFGNTIKVPTPLCARQODE 1888
Db 314 ILMGYRFPVPLTLDEGFYEVNDYNSFHETVETSTSLSKAEIAE 356

RESULT 8
US-10-995-561-967
; Sequence 967, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 967
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-967

Query Match          8.2%; Score 799; DB 6; Length 268;
Best Local Similarity 56.3%; Pred. No. 1.5e-51;
Matches 151; Conservative 56; Mismatches 61; Indels 0; Gaps 0;

Qy 1264 LKESYTEKAPWQIQETAPSSWPQVGRVERFNCLRYREDLDFVLAHINVTINGEKV 1323
Db 1 MDVYMTPEKEAPWRLPTCAQPPWPGQIQIFRDFGLRYRRELPLAVQGVSFKIHAGKVV 60
Qy 1324 GIVGRTGAGKSLTGLFRINESAGEELIIDGINAKIGLDLRKTIITIPQDPVLPSGS 1383
Db 61 GIVGRTGAGKSSLASGLRLQEAAGGIWIDGVPLAHGHLTKLRISITIIIPQDPLPFGS 120
Qy 1384 LRMDLPSQXSDEEVTSLSLAHKDPVSAIPDLTDEHCAEGENLTVGQROVLCLARA 1443
Db 121 LRMDLDQHSDELAVALAETVQAKALVASTPGLQYKCKADRGEDLSVGQKQLCLARA 180
Qy 1444 LIRKTKIIVLDEATAAVDLJETDDLIOSTIRQFEDCTVLTIAHRIINTMDYTRVIVLDKG 1503
Db 1444 LIRKTKIIVLDEATAAVDLJETDDLIOSTIRQFEDCTVLTIAHRIINTMDYTRVIVLDKG 1503
```

Db 181 LNRKQTLILDEATAVDPGTELOQMALGSMFAQCTVLLIAHRLASVMDCAVLAWMKG 240
QY 1504 EIOEYGAUSDLLQORGLFYEMAKDAGLV 1531
Db 241 QVAESSGPAQLAQKGLFYRLAQESSGLV 268

RESULT 9

US-10-613-744-13
; Sequence 13, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: NO PCT/US99/06307
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentln Ver. 2.1
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-613-744-13

Query Match 8.1%; Score 786.5; DB 6; Length 391;
Best Local Similarity 46.2%; Pred. NO. 2.2e-50;

Matches 150; Conservative 75; Mismatches 89; Indels 11; Gaps 6;

QY 1564 RTERRARFVSKGNCNVAKHNIREQRFL--QDVFTLLVDLKKPHLLIFTMSFLCSWL 1621
Db 35 RSNQ-RARLVSKSGRCNIERGVNDAQSRPIPFVDIWTTVLDKKRYQMTVFITAFGLSMF 93
QY 1622 LFPAMWMLFPAHGD---APGGINVPCYTSIHSFSSAFLEISVQVITIGGCGRVTE 1677
Db 94 LFGILWVVAVYVHVDLPEFPYPPDNR--PCVENINGMTSAFLFSLFTQVITIGYGFRTVE 151
QY 1678 ECPALILILVONVGMNAINMIGCFMKTAAQHRRAETLIPSKHAVITLRRHRLCFML 1737
Db 152 QCATALFLIPQSLIGVINSFPGALIAKISRKKAKTITTSKNAVISKRGKCLLI 211
QY 1738 RVGDLKSKMISATIHQVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAPLIYHV 1797
Db 212 RVANLRKSLIGSHIYVGLKLTITTPBGETIILDQTNINPVNAGNMLFFISLTIYH 271
QY 1798 IDSNSPIYDLAPSDLHHODLEIIVILEGVETTGITTOARTSYLADEILMGORFPIVA 1857
Db 272 IDHNSPFFHMAAETL--SQDFELVVPFDGTVESTATCQVTSYVPEVLMGVRFYIVS 330
QY 1858 E-EDGRYSVYSKFGNTIKYPTPLC 1881
Db 331 KTEGKYVDVFNHFKTVEVETPHC 355

RESULT 10

US-10-793-626-326
; Sequence 326, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US

CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 326
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-326

Query Match 5.3%; Score 514.5; DB 6; Length 548;
Best Local Similarity 27.1%; Pred. No. 3.5e-30;

Matches 157; Conservative 110; Mismatches 211; Indels 101; Gaps 16;

QY 1002 DPLVNGQERTKRLSYGALGISQGIADVGYMAVSIIGLISRCILHVDLHSLTASPM 1061
Db 12 DGVIN--NHSLTNOEKFSHGLVAIGIALP-----IFLIVRPPI 47
QY 1062 SFPE---RTPSGNLVVRFSKEL-----DTVDSMIPEYIK-----MEWG 1096
Db 48 EPIRQYLAQMTSKILYDIRKQLYNHLQALSVFANNQGVYISRYINDVEQTKDFILT 107
QY 1097 SLFNVIGACIVILLAPPI-----AAIIPPLGL--IYFVQRFVYASSRQLRLE 1144
Db 108 GLNMIMWDCITIIILASIMFPLDVKLTFLAIFIPFYILTVYFFGRL-----RKLIRVR 162
QY 1145 SVSRSPYSHFNETLGVSTIRAPESQERFIHQSDLKVDENQAKAYYSIYANRLAVRLE 1204
Db 163 SQALAEVQGFLEHRYVQMSVYKSFALIDNEAKNF--NNKQFLQRAFOHTRWNAVSA 219
QY 1205 CV-----GNCIVLEALFAVISRHSLSAGLVSVSYQVTTYLMVLVMSSEMETNI 1258
Db 220 AINTVTDLGPILYIIGVSYLAIT--GSTVGTTLAIFYQLGFLRLVSSFTTLIQSF 278
QY 1259 VAVERLKEVSETEKE-----APWQIQTAPPSWPQVGRVFRNYCLAREDDLPVLR 1311
Db 279 ASMDRVQLMDEDDYDKNGIQAQPIKSK-----QIDDKVSPFYNEKEKVLH 328
QY 1312 HINVTINGEKVQIVERTGAGKSSLTGLFRINESAGEIITIDGINAKIAGLDLRKIT 1371
Db 329 DINTLNKGETVAFVGGSGKSTLNLIPRFYDVTOGEIILIDHNVKDFLTGSLRNOIG 388
QY 1372 IIPQDPVLFSGSLRML---DPSQYSDEEVTWSLELAHKDFVSLPDKLDHECAAGE 1428
Db 389 LVQDDNLTFSQDYKENILGLRP--DATDDEVEBAKKAANAHDTISLNPNGYDTEVGERGV 446
QY 1429 NLSVGQROLVCLARALLRKTILVLDATVAVDLETTDLIOSTIRQFEDCTVLTIAHRL 1488
Db 447 KLSGGQKQRISTARIFLNNPVLIDBATSALDBESAIIOELADVLVSKORTLTIYVHRL 506
QY 1489 NITMDYTRVIVLDKSGISOEYGAUSDLLQORGLF---YSM 1524
Db 507 SITTHADRIYVWENGRIYVETGTHQQLINKGAYEHLYSI 545

RESULT 11

US-10-467-657-1346
; Sequence 1346, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZAZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657

```
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1346
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1346

Query Match      5.1%; Score 495.5; DB 6; Length 622;
Best Local Similarity 28.7%; Pred. No. 1e-28;
Matches 154; Conservative 106; Mismatches 213; Indels 63; Gaps 13;

QY 1022 LGISGIAVFGYSMAVSIIGGLARCLHVDLHSLIRSPMSFPFERTPSGNVNFPSKELD 1081
DB 108 LVVIRGICRFSTYTLMTWVSMTTSKIKDMFAKMLTSSRYHQTPTSGTVLMMNLITE 167
QY 1082 TVDSMIEVIMKFMGSLFNVIAGACIVLL---ATPIAIIIPPLGLIYFVFORFYVASS 1137
DB 168 QSVNASDIFTVLRDNTIVGLTIVLYLNMQLSLIYVLMFPLSL---SRY---R 220
QY 1138 RQLKRLSVSRSPVYSHFN---ETLLGVSVIRAFEEGRFTHQSDL-----KVDE 1184
DB 221 DRLKHVIDSOKSIGTMMNVIAETHQGRVVKLFNGQAQANRPDAVNRITVRLSKITQ 280
QY 1185 NOKAYYP--SIVARNMLAVRLECVGNCIVLPALPAVISRHSLSAGLVGSYSLOYTT 1242
DB 281 ATAAHSPSELSIAIALAV-----VITIALQSQNGYTTTIGEFMAPIVA--MLQMYA 330
QY 1243 YLNLVMSSEMETNIVAVRLKEYSETEKE-----APWQIOETAPPSWPQVGRVER 1296
DB 331 PIKSLANISIPMQTMFLAAGVCAFLDTPPEODKGTALPQRYE-----GRISFR 379
QY 1297 NYCARYEDDLYVRHINVTINGEKVGIVRTGAGKSSLTGLFRLNESABGEIITDGI 1356
DB 380 NVDEYRSYSGIKALDNFNLDRQGERVALVGRSGSKSTVNLPRFEYPAGNICIDGI 439
QY 1357 NIAKIGHDLRFKTIITIPQDPVLFSGLRMULDPSQYS-----DEEYTSLELAHKDF 1411
DB 440 DIADIKDCLRAQALVSGDFLEPDTLPFENV---KRSRPDAGEAELVSLQAAINLSL 495
QY 1412 VSALPDKLDHBCAEGENLAVGROQLVCLARALLRKTKILVDEATAVADLETDDLQST 1471
DB 496 IDASPLGLHQPIGSGNSLSCGGRQQRVALARAILLKADAILLDDETSALDNESEELVQA 555
QY 1472 IRTQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILQORGLFYSMAD 1527
DB 556 LERLMENRTGIIVAHRLTTVESADRIIVMDGKIIIEQTHDQLMFQNG--YVTMLRN 610

RESULT 12
US-11-082-389-148
; Sequence 148, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroeger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zeider, Oskar
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131CPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
```

```
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 148
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-082-389-148

Query Match      4.8%; Score 468.5; DB 7; Length 549;
Best Local Similarity 26.9%; Pred. No. 8.3e-27;
Matches 141; Conservative 110; Mismatches 201; Indels 73; Gaps 18;

QY 1031 FGYSMAVSIIGGLIARCLHVDLHSLIRSPMSFPER-----TPSGNLVNFPSKELD 1084
DB 50 FGSKLSMEVGR-----DLRSALRGKVVNFSEEREMGQGAAS--LITNINDVOOQ 98
QY 1085 SMIEVIMKFMGSLFNVIAGACI-----VILATPIAIIIPPLGLIYFVFORFY 1133
DB 99 MLVQMTSTLMSAPMLAIGIIMAVRODLGLSLMVMVSIPLIIV--ALIIYRMVPLF 155
QY 1134 VASRQLKRLSVSRSPVYSHFNETLLGVSVIRAF---EEQERFIHOS-----DLKVD 1183
DB 156 QTMOKRIDRIHQIIR-----EOLTGIRVIRAFVBEVDERERFTTASKVDADIGVETG 207
QY 1184 ENOKAYYSIVANMLAVRLECVGNCIVLPALPAVISRHSLSAGLVGSYSLOYTT 1243
DB 208 NMLAMFPAVW---LIMTSLAV---VIMGAFQVSGET-QIGTLFAPLQYIMQILMG 259
QY 1244 LNLVMSSEMETNIVAVRLKEYSETEKEAPWQIOET--APPSSWPQVGRVERFYCYRY 1302
DB 260 VMMAAFHFWVWPRAAVSADRIGEVLETPSY--QAPETPAQPT--SAGEIVFNNAFAY 315
QY 1303 REDIDPVLHINVTINGEKVGIVRTGAGKSSLTGLF--RINESABGEIITDGINAKI 1361
DB 316 PGADDPVLNNVSPFAVSGSTTAIIIGTSGGKTTL--IGLVPRLPVTEGDVTVDDVDAEF 374
QY 1362 GLHDLRFKTIITIPQDPVLFSGLRMULDPSQYSD-----EWYTSLELAHKDFVSLP 1416
DB 375 EPLKMDRIGLVPPKSLFSGTIANSL---RYGNEDATETQLWOLALIAQADAPVREMP 430
QY 1417 DKLDECAEGENLAVGROQLVCLARALLRKTKILVDEATAVADLETDDLQSTIRTOF 1476
DB 431 EGLDSEIAQSGTIVSGGRQLALARALLKPELIYIPDDSSALDVSIDALRRALSTNL 490
QY 1477 EDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILQORGLF 1521
DB 491 PDATKLIYAGRVSTIRADQIVVDNGBGVIGITHTMLNMGTY 535

RESULT 13
US-10-995-561-575
; Sequence 575, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
```

```

; FILE REFERENCE: CLO015C9
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FASTSEQ for Windows Version 4.0.
; SEQ ID NO 575
; LENGTH: 701
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-995-561-575

```

Query Match	4.3%	Score 419.5;	DB 6;	Length 701;
Best Local Similarity	25.2%	Pred. No. 4.8e-23;		
Matches 149;	Conservative 130;	Mismatches 234;	Indels 79;	Gaps 21;

```

OY 977 |S|F|F|M-|C|H|N|V|A|S|A|N|N|N|L|S|T|D|D|P|I|N|V|N|Q|O|E|H|K|N|L|S|V|G|A|L|G|I|O|G|I|A|V|P|G|S 103
D| 140 |L|S|F|L|V|L|V|L|S|S|G|E|W|A|I|P|F|F|G|R|L|D|I|L|Q|G|S|A|D|T|R|N|L|T|S|I|L|T|A|S|A|V|E|F| 196
OY 1035 |M|A|V|S|I|G-|G|I|L|A|S|R|C|H|V-|D|L|H|S|I|L|S|P|M|S|F|E|R|T|S|C|N|L|N|R|F|S|K|E|L|D|V|D|S|I 1087
D| 197 ----|V|G|D|I|Y|N|N|T|G|H|V|H|S|I|O|E|V|F|G|A|V|L|Q|E|T|E|F|P|O|N|Q|T|N|S|K|V|T|E|D|T|S|L 252
OY 1088 |E|V|I|R|E|K|S|L|F|N|V|I|G|A|C|V-|I|L|----|A|P|I|A|I|I|P|P|G|L|V|F|V|O|P|F|V|A|S|R|O|L|K 1141
D| 253 |S|E|N|S|I|F|F|M|Y|L-|V|R|G|L|C|I|G|I|M|I|G|S|V|L|T|V|T|I|T|P|L|F|L|P|K|R|K|V|K|Y|Q|L|E|V|Q|V|R 310
OY 1142 |R|L|E|V|S|R|P|V|Y|S|H|R|E|T|L|G|S|V|R|A|F|E|O|R|F|I|H|O|S|L|K|Y|D|----|N|O|K-|A|Y|P|S|V 1194
D| 311 --|E|S|I|A|S|Q|V|A-|I|E|A|S|A|P|T|V|A|S|F|A|N|E|G|E|A|Q|F|R|E|K|I|Q|E|I|K|T|I|N|Q|K|A|V|A|----|A 362
OY 1195 |A|N|R|W-|L|A|R|E|C|V|N|C|Y|L|F|A|L|F|A|V|I|S|R|H|S|I|S|A|G|L|V|G|L|S|Y|S|I|Q|T|V|Y|L|N|W 1246
D| 363 |V|N|S|T|T|S|I|S|G|L|K|V|G|I|Y|I|G|Q|L|V|T|S|A-|-----|V|S|G|N|L|V|T|P|Y|Q|M|O|F|Q|A|V|E|V 413
OY 1247 |L|V|R|N|S|E|K|E|T|N|I|V|A|V|R|L|K|E|Y|S|E|T|E|K|A|P|W|O|I|Q|E|T|A|P|S|S-|W|P|O|Y|R|V|E|F|R|N|Y|C|U|R 1302
D| 414 |L|S|I|P|R|V|Q|A|K|V|S|E|K|I|F|E|Y|L|D|R|T|R-|-----|C|R|P|S|G|L|T|P|L|H|R|G|L|V|Q|O|V|S|P|A|Y 465
OY 1303 --|R|E|D|D|P|R|V|L|R|H|I|N|V|T|I|N|G|R|E|K|V|G|I|V|G|T|G|A|G|K|S|L|T|I|G|L|R|I|N|S|A|G|E|I|I|D|I|G|I|N|A|K 136C
D| 466 |P|R|R|D|V-|L|V|O|G|L|F|T|R|P|R|G|E|V|T|A|L|V|G|N|G|S|G|K|S|V|A|L|L|N|L|V|O|P|T|G|O|L|L|D|G|K|P|R|Q 524
OY 1361 |I|G|L|H|D|R|E|K|T|I|I|P|O|D|P|V|L|F|S|G|S|I|R|A|N|L-|D|P|F|Q|S|D|E|E|V|W|T|S|L|E|L|A|H|D|P|V|S|A|L|P|D|X 141B
D| 525 |Y|E|H|Y|I|R|Q|V|A|A|V|Q|E|P|O|V|P|R|G|S|I|Q|E|N|I|A|Y|G|L|T|O|K|P|T|H|E|E|I|T|A|A|V|S|G|A|S|F|I|S|G|L|P|O|G 584
OY 1419 |L|D|H|E|C|A|G|E|N|L|S|V|G|R|Q|V|C|I|A|R|A|L|K|R|K|I|L|V|D|E|A|T|A|V|D-|L|E|T|D|L|I|O|S|T|I|R|T 1474
D| 585 |Y|D|V|E|I|D|E|A|G|S|O|L|S|G|G|R|Q|A|V|A|L|R|A|R|I|R|K|P|C|V|L|I|D|D|A|T|S|A|L|D|A|N|S|Q|V|O|B|L|Y|E|S-|P 642
OY 1475 |Q|F|E|C|T|V|T|I|A|R|A|N|T|I|N|D|Y|R|V|V|D|B|G|E|I|Q|E|T|A|B|S|D|L|Q|O|G|L|F|I|S|A|K 1526
D| 643 |B|R|Y|S|R|V|L|I|Q|H|L|S|V|R|Q|A|D|H|I|L|F|L|E|G|A|R|E|G|G|H|Q|O|L|E|K|G|C|Y|W|A|V|O 694

RESULT 14
US-10-995-561-574
; Sequence 574, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 574
; LENGTH: 808
; TYPE: PR1

```

ORGANISM: Homo sapiens
US-10-995-561-574

Query Match	4.3%	Score 419.5	DB 6	Length 808
Best Local Similarity	25.2%	Pred. No. 5.9e-23		
Matches 149	Conservative 130	Mismatches 234	Indels 79	Gaps 21

Oy 977 LSIPLFM--CNHVSALSNWYLSLMTDPIVNGTQEHKXVLSYVGALGISOGIAVFGYS 103
 Db 247 LSLPLVLVLSLSLGMALPFPTGSLTBLMLODGSADFTFRLTLMSLITLISAVLFE-- 303
 Oy 1035 MAVSIG-GILASRLCY-----DLHSLIRSMPSPFEERTSGNVLVNFPSKELDTVDSMI 1087
 Db 304 ---VGDIYNNNTGHHSHSLQGEVFGAVLQETEFPOQNTQMSRVTEDTSTLSDSL 359
 Oy 1088 PEVIKMGMSLFNIGACIV-IL------APPIAIIIPELGILYFPQRYVASSRDLK 1144
 Db 360 SENISLPLMWY--VRGCLGIGIMMGSVSLMTVTLITPLFLFKPKYKXWYQLLEVOVR 417
 Oy 11442 RLESVRSRPVYSHNETLLAGVSVIRAFEEORFIHOSDLKYDE-----NOK--AVYPSIV 1194
 Db 418 --BELASSQVA--IEHLSMPTVRSPANEGEAGQKREKQIEIKTLNQKRAVY----A 469
 Oy 1195 ANRW-----LAVRLECGNCIVLPALFAVISRSLASGLVGSYSLOVTTYLNM 1246
 Db 470 VNSMTTISIGMLTKVGLIYIGGQLVTGSA-----VSSGNIJVFVLXOMQFOGAVEV 520
 Oy 1247 LVRRSMETNIVAVEBLKEYSETEKEAPWQIOETAPSS----WPQGRFEPNYCLRY 1302
 Db 521 LLSITPRVQKAVSGSEKIFETLDTTPR-----CPBSGLITPLHEGLVQFOVDSFAY 572
 Oy 1303 --REDLDFVLRIHNVITNGGEKVGIVSGRTGAKSSLTGLFRINESAGEIITIDGINIAK 1366
 Db 573 PNRPDV-LVLQGLFTFLRPGFVTLVGPNGSGKSTVALLONTLYOPTGGQLLDOCKPLPQ 631
 Oy 1361 IGLHDLAFKTIITIQDVLVPSGSLRML--DPFSQYSDDEVWTSLELAHLKDPVSALPDK 1418
 Db 632 YEHRYLIRQVAAVQOEPOVFERSLQENIAYGLTQKPTMEETLAAVXGASHFISGLPQG 691
 Oy 1419 LDHCEACGEGENLSVGORQLVCLAPALRKTKIIVLDEATAVD---LETBDLIQSTIRT 1474
 Db 692 YDTEVDEBAGSQLSGGQRQVAAALAPALRKPCVLLIDTATSLMANSQQLVQESLYES--P 749
 Oy 1475 QFEDCTVLTIAHRLINTMDYTRVIVLDKGEIOEYAGBSDLIQRGLEFYSMAK 1526
 Db 750 ERYRSRVLLITQHLISLVEQADHILFEGAIRGEGTHIQOLMEKKGCYAWAQ 801
 RESULT 15
 US-11-105-268-53
 ; Sequence 53, Application US/11105268
 ; Publication No. US20050260204A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allan, Christian
 ; TITLE OF INVENTION: ANTI-IL-9 ANTIBODY FORMULATIONS AND USES THEREOF
 ; FILE REFERENCE: 10271-126-999
 ; CURRENT APPLICATION NUMBER: US/11/105, 268
 ; PRIOR FILING DATE: 2005-04-12
 ; PRIOR APPLICATION NUMBER: 60/561,845
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 53
 ; LENGTH: 808
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-105-268-53
 Query Match 4.3%; Score 419.5; DB 7; Length 808;
 Best Local Similarity 25.2%; Prod. No. 5.9e-23;
 Matches 149; Conservative 130; Mismatches 234; Indels 79; Gaps 21;
 Oy 977 LSIPLFM--CNHVSALSNWYLSLMTDPIVNGTQEHKXVLSYVGALGISOGIAVFGYS 1034

977 LSIFLFM--CNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYGALGISQGIAVFGYS 1034

Db 247 LSLFLVLVLSLSLGMALPFTGRLLDWILODSADTFTRNLTLMSILTLIASAVLEF--- 303
QY 1035 MAVSIG-GILASRCLHV-----DLHSILRSPMSFPERTPSGNLVNRFSKELDTVDSMI 1087
Db 304 ----VGDCIYNNMTGHHVSHLOGEVFGAIVROETEFPOONQGNIMSRVTEDTSTLSDSL 359
QY 1088 PEVIKMFWSLFNVIGACIV-ILL-----ATPIAIIIPPLGLIYFVQRFYVASSRQLK 1141
Db 360 SENISLFLMWL--VKGCLGLGIMLWGSVSLTNVTLITLPLFLPKVKGVQYLLLEVQVR 417
QY 1142 RLESYSRSPVYSHFRETLLGVSVAIFAEQERFIHQSLKYDE---NOK--AYPSIV 1194
Db 418 --ESLAKSSQVA--IEALSAMPYVRSFANEBGEAQFREKLOEIKTLNQKEAVAY---A 469
QY 1195 ANRW-----LAVRELCVGCIVLFAALFAVISRHSISAGLVGLSVSYSLQVTTYLNM 1246
Db 470 VNSWTTISGMLKKVGIYIGQQLVTSQA-----VSSGNLVTFVLYQMOTQAVEV 520
QY 1247 LVRRSSEMETNIIVAVERLKEYSETEKEAPWQIQETAPPS---WPOVGRVEFRNYCLRY 1302
Db 521 LLSIYPRVQKAVGSSEKIFEVYLDRTPR-----CPPSGLLTPLHLGLVQFDVSPAY 572
QY 1303 --REDLDBVLAHINVTINGEKNVIGRTGACKSLTGLFRINESARGEIITDGINIAK 1360
Db 573 PNRPDV-LVLOGLTFTLRPGEVTLVGPNGSGKSTVAALLOVLXOPTGQQLLDGKPLPQ 631
QY 1361 IGLHDLRFKTIITIPQDPVLFSGSLRMNL--DPFSQYSDGEVWTSLELAHLKDFVSALPDK 1418
Db 632 YEHRYLHRQVANAAGDEPOVFGKSLQENIAYGLTQKPTMEBITAANA VKSGANHSFISGLPQG 691
QY 1419 LDHRCABEGENLSVGQQLVCLARALRKTKILVLDEATAVD---LETDDLQSTIRT 1474
Db 692 YDTEVDEAGSQLSGGQQAVALARALIRKPCVLILDDATSLDANSQLOVEQLYES--P 749
QY 1475 QFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYCAPSDLLQQRGLFYSMAX 1526
Db 750 ERSRSVLLITQHLSLVEQADHILFLEGGAIREGGTHOQLMEKGCYMAWVQ 801

Search completed: December 15, 2005, 15:50:09
Job time : 11.1294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:27:22 ; Search time 173.071 Seconds
(without alignments)
4800.712 Million cell updates/sec

Title: US-10-665-283-4

Perfect score: 9734
Sequence: 1 MALRGFCSAGSDPLMDMNV.....NTIKVPPLCTARQLDSDRS 1891

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	9734	100.0	1891	9 ADY86941	Ady86941 Human MRP
2	9734	100.0	1927	9 ADY86943	Ady86943 MRP1-mous
3	9734	100.0	1927	9 ADY86938	Ady86938 Human MRP
4	9704	99.7	1947	9 ADY86945	Ady86945 Human MRP
5	7860	80.7	1531	2 AAW57486	Aaw57486 Human MRP
6	7860	80.7	1531	2 AAW74471	Aaw74471 Human mul
7	7860	80.7	1531	2 AAW98994	Aaw98994 Human mul
8	7860	80.7	1531	3 AAY55799	Aay55799 Human mul
9	7860	80.7	1531	3 AAY78873	Aay78873 Multidrug
10	7860	80.7	1531	5 ABG61810	Abg61810 Prostate
11	7860	80.7	1531	6 ABM35012	Abm35012 Cancer ba
12	7860	80.7	1531	6 ADB20865	Adb20865 MRP1 base
13	7860	80.7	1531	7 ADB87954	Adb87954 Human UGT
14	7860	80.7	1531	7 ADB96937	Adb96937 Human MDR
15	7860	80.7	1531	7 ADB92128	Adb92128 Human MDR
16	7860	80.7	1531	7 ADD44764	Add44764 Human pro
17	7860	80.7	1531	8 ADN97111	Adn97111 MRP1 prot
18	7860	80.7	1531	8 ADU06426	Adu06426 Novel bro
19	7860	80.7	1531	9 ADX06202	Adx06202 Cyclin-de
20	7856	80.7	1530	9 AADN95929	Adn95929 Human BEC
21	7849	80.6	1531	2 AAR54928	Aar54928 Multidrug
22	7849	80.6	1531	2 AAR93153	Aar93153 Multi-dru
23	7849	80.6	1531	2 AAW57485	Aaw57485 Human mul
24	7849	80.6	1531	2 AAW74470	Aaw74470 Human mul

25	7849	80.6	1531	2 AAW98993	Aaw98993 Human mul
26	7849	80.6	1531	3 AAY55798	Aay55798 Human mul
27	7849	80.6	1531	3 AAY78872	Aay78872 Human mul
28	7849	80.6	1531	3 AAB03582	Aab03582 Multidrug
29	7763	79.8	1515	8 ADU24094	Adu24094 Human Cys
30	7610	78.2	1489	2 AAR66952	Aar66952 Multi-dru
31	7587.5	77.9	1482	8 ADN03902	Adn03902 Antipsori
32	7587.5	77.9	1482	8 ADP23422	Adp23422 PRO poly
33	7572	77.8	1481	5 ABP52108	Abp52108 Homo bapi
34	7560.5	77.7	1482	2 AAR66953	Aar66953 Multi-dru
35	7535	77.4	1475	9 ADX06206	Adx06206 Cyclin-de
36	7510.5	77.2	1472	9 ADX06204	Adx06204 Cyclin-de
37	7500.5	77.2	1472	2 AAR66954	Aar66954 Multi-dru
38	7230.5	74.3	1530	9 ADP56611	Adp56611 Bovine MR
39	7195.5	73.9	1416	9 ADX06208	Adx06208 Cyclin-de
40	7194	73.9	1417	2 AAR66955	Aar66955 Multi-dru
41	7037	72.3	1388	7 ADJ69912	Adj69912 Human hea
42	7002.5	71.9	1528	2 AAW57487	Aaw57487 Murine mu
43	7002.5	71.9	1528	2 AAW74472	Aaw74472 Mouse mul
44	7002.5	71.9	1528	2 AAW98995	Aaw98995 Mouse mul
45	7002.5	71.9	1528	3 AAY55800	Aay55800 Murine mu

ALIGNMENTS

RESULT 1
ID ADY86941 standard; protein; 1891 AA.
XX
AC ADY86941;
XX
DT 02-JUN-2005 (first entry)
XX
DE Human MRP1-Kir6.2 mutant delc36 fusion protein, SEQ ID NO: 4.
XX
KW Ionophore; biosensor; drug screening; diagnostic;
KW microorganism detection; potassium channel; fusion protein;
KW multidrug resistance protein 1; MRP1; Kir6.2; mulein.
XX
OS Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX
FH Key
FH Region 1..1531
FT /note= "Multidrug resistance protein 1 (MRP1)"
FT Region 1532..1537
FT /note= "Hexaglycine spacer"
FT Region 1538..1891
FT /note= "Kir6.2 mutant delc36 protein"
XX
US2005063989-A1.
XX
PD 24-MAR-2005.
XX
PF 22-SEP-2003; 2003US-00665283.
XX
PR 22-SEP-2003; 2003US-00665283.
XX
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX
PI Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
XX WPI; 2005-252611/26.
XX
PT New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
XX contaminant/pollutant in a sample.
XX
PS Claim 25; SEQ ID NO 4; 78pp; English.
XX
CC The present invention relates to a hybrid protein consisting essentially

CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for the screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multiling resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 mutant detac36 protein.
XX

SQ Sequence 1891 AA;

Query Match 100.0%; Score 9734; DB 9; Length 1891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLMDNNTWNTSNPDFTKCFQNTVLAWVPCFYLMACPFFYLYLSRH 60
Db 1 MALRGFCSDGSDPLMDNNTWNTSNPDFTKCFQNTVLAWVPCFYLMACPFFYLYLSRH 60
QY 61 DRGYIQMTPLNKKTALGFLIMVCMADLPYSFWRSGIFLAPVFLVSPITLGIITLLA 120
Db 61 DRGYIQMTPLNKKTALGFLIMVCMADLPYSFWRSGIFLAPVFLVSPITLGIITLLA 120
QY 121 TELIOLERRKGVSSGIMLTFWLVALVCAALIRSKIMTALKEDAQVDLFRDITFFYYS 180
Db 121 TELIOLERRKGVSSGIMLTFWLVALVCAALIRSKIMTALKEDAQVDLFRDITFFYYS 180
QY 181 LLLIQLVLSGSDSPLESETIHDNPPCRESSASFLSRITFMWITGLIVRGYROPLEGSD 240
Db 181 LLLIQLVLSGSDSPLESETIHDNPPCRESSASFLSRITFMWITGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSQVVPVLVKNKKCECAKTRKOPVKKVYSSKDPACPKESSKVDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVLVKNKKCECAKTRKOPVKKVYSSKDPACPKESSKVDANEVEAL 300
QY 301 IVKSPKEMNSLKFVLVYKTFGRPYFLMSFFPKAIHDLMPSPGPIILKILFVNDTKAPD 360
Db 301 IVKSPKEMNSLKFVLVYKTFGRPYFLMSFFPKAIHDLMPSPGPIILKILFVNDTKAPD 360
QY 361 WQGFYFVTLFVTACIQTLVLAHOYFHCIFVSGMRITKAVI GA VYKALVITNSARKSSTV 420
Db 361 WQGFYFVTLFVTACIQTLVLAHOYFHCIFVSGMRITKAVI GA VYKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMTWSAPLQVILALYLLMLNGSPSLAGVAVMLVMEVN 480
Db 421 GEIVNLSVDAQRFMDLATYINMTWSAPLQVILALYLLMLNGSPSLAGVAVMLVMEVN 480
QY 481 AVMAKTKTYOVAMHKSNDRIKLMNELNGIKVLKLYAMELAFQDKTLAROEELKYLK 540
Db 481 AVMAKTKTYOVAMHKSNDRIKLMNELNGIKVLKLYAMELAFQDKTLAROEELKYLK 540
QY 541 KSAVLSAVGFTWCTPFLVALCTPAVYVITDENNIIDLAQTA FVSALFNILRPFLNLTLP 600
Db 541 KSAVLSAVGFTWCTPFLVALCTPAVYVITDENNIIDLAQTA FVSALFNILRPFLNLTLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPYKDGCTNSITVRNATFTWASDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPYKDGCTNSITVRNATFTWASDPT 660
QY 661 LINGITPSPGALVA VVQVCGKSSLSALLAENDKKEGHVATIGSAVYVPOQAMIOND 720
Db 661 LINGITPSPGALVA VVQVCGKSSLSALLAENDKKEGHVATIGSAVYVPOQAMIOND 720
QY 721 SLRENIIFGCOLLEPYRSVIOACALPDLEILPSGRTIEGKGVNLSGGOKORVSLAR 780
Db 721 SLRENIIFGCOLLEPYRSVIOACALPDLEILPSGRTIEGKGVNLSGGOKORVSLAR 780
QY 781 AVYSNADIIYLPDPLSAVDAAVGHKIPENVI GPKGMLNKTRILVTTHSMSTLPQVDVILV 840
Db 781 AVYSNADIIYLPDPLSAVDAAVGHKIPENVI GPKGMLNKTRILVTTHSMSTLPQVDVILV 840
QY 841 MSGGKISEMSGYOEILARDGAFAEFLRTYASTEOQDAEENGVTGVS GPKGAKOMENGM 900
Db 841 MSGGKISEMSGYOEILARDGAFAEFLRTYASTEOQDAEENGVTGVS GPKGAKOMENGM 900

Db 841 MSGGKISEMSGYOEILARDGAFAEFLRTYASTEOQDAEENGVTGVS GPKGAKOMENGM 900
QY 901 LVYDSAGKOLQROUSSSSSYSGDISRHNNSTAELOKAEKEETWKLMEADKAOTGVYKL 960
Db 901 LVYDSAGKOLQROUSSSSSYSGDISRHNNSTAELOKAEKEETWKLMEADKAOTGVYKL 960
QY 961 SVYWDYKAIQGLFISPLSIFLFWCNHVSALASNYWLSLWTDPIVNGTOEHTKRLSYYG 1020
Db 961 SVYWDYKAIQGLFISPLSIFLFWCNHVSALASNYWLSLWTDPIVNGTOEHTKRLSYYG 1020
QY 1021 ALGISOGIAVGVSMVAISIGIILASRCLHVDLHSILRSPSPFEETPSGNLVNRSKEL 1080
Db 1021 ALGISOGIAVGVSMVAISIGIILASRCLHVDLHSILRSPSPFEETPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPEVILKMFMSGLFNVIGACIYLATPAAIIPPLGIYEFVQGFYVASSROL 1140
Db 1081 DTVDSMIPEVILKMFMSGLFNVIGACIYLATPAAIIPPLGIYEFVQGFYVASSROL 1140
QY 1141 KRLSVSRSPYSHFNFTLLGVSYIRAFEEQERFIHOSDLKVDENQAKYPSIVANRWLA 1200
Db 1141 KRLSVSRSPYSHFNFTLLGVSYIRAFEEQERFIHOSDLKVDENQAKYPSIVANRWLA 1200
QY 1201 VRLCEVNCIYLFALFAVILSRHLSAGLVGLSVSYLOVTTYLNMVLRNSESMEETIVA 1260
Db 1201 VRLCEVNCIYLFALFAVILSRHLSAGLVGLSVSYLOVTTYLNMVLRNSESMEETIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPOVGRVFRNYCLRYREDLPVLJHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPSSWPOVGRVFRNYCLRYREDLPVLJHINVTINGG 1320
QY 1321 EKVGIVGRTGAKSSLTGLFRIINESABGIIIDGINIAKIGLHDLAPKTIITIQDPLVF 1380
Db 1321 EKVGIVGRTGAKSSLTGLFRIINESABGIIIDGINIAKIGLHDLAPKTIITIQDPLVF 1380
QY 1381 SCSLRMNLDPSPQSDSEBWTSLBLAKDPVSAIPKLDHECAGEGNSVSGOROLVCL 1440
Db 1381 SCSLRMNLDPSPQSDSEBWTSLBLAKDPVSAIPKLDHECAGEGNSVSGOROLVCL 1440
QY 1441 ARALIRKTKILVDEATVAVDLETDLIIOSTIRTOFEDCTVLTIAHRLNTIMDTTRYVL 1500
Db 1441 ARALIRKTKILVDEATVAVDLETDLIIOSTIRTOFEDCTVLTIAHRLNTIMDTTRYVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYMAKDAGLVGGGGGMLSRKGIIPESYVLTRLAEDPAE 1560
Db 1501 DKGEIOEYGAPSDLLQORGLFYMAKDAGLVGGGGGMLSRKGIIPESYVLTRLAEDPAE 1560
QY 1561 PRYTRRRRARFVSKKNCNVAAHNIREOGRFLODVFETTLVDLKWPHTLILFTMSPLCSW 1620
Db 1561 PRYTRRRRARFVSKKNCNVAAHNIREOGRFLODVFETTLVDLKWPHTLILFTMSPLCSW 1620
QY 1621 LLFAMVWMLIAFAHGDAPGEGTVPCVTSIHSFSAFLPSIEVOVITGFGRWVTEBCP 1680
Db 1621 LLFAMVWMLIAFAHGDAPGEGTVPCVTSIHSFSAFLPSIEVOVITGFGRWVTEBCP 1680
QY 1681 LAIILIVONIVGMINAIVMLGCIEMKTAQHARRAFTLISKBAVITLHGRCLCFMILRVG 1740
Db 1681 LAIILIVONIVGMINAIVMLGCIEMKTAQHARRAFTLISKBAVITLHGRCLCFMILRVG 1740
QY 1741 DLRSMTIISATIHQVVRKTTSPGEVAPPLHQVDIPENGVGGNGITFLVAPLIIYHYIDS 1800
Db 1741 DLRSMTIISATIHQVVRKTTSPGEVAPPLHQVDIPENGVGGNGITFLVAPLIIYHYIDS 1800
QY 1801 NSPLYDLAPSDLIHHODLEIIVILEGVEETGTTQARTSYLADBIIMGGRFVPIVAEED 1860
Db 1801 NSPLYDLAPSDLIHHODLEIIVILEGVEETGTTQARTSYLADBIIMGGRFVPIVAEED 1860
QY 1861 GRYSVDYSKFGNTIKVPTPLCTARQDLDEDS 1891
Db 1861 GRYSVDYSKFGNTIKVPTPLCTARQDLDEDS 1891

RESULT 2
ADY86943

ID ADY86943 standard; protein; 1927 AA.
XX
AC ADY86943;
XX
DT 02-JUN-2005 (first entry)
XX
DE MRP1-mouse K1r6.2 mutant (KR370AA) fusion protein, SEQ ID NO: 6.
XX
KW Ionophore; biosensor; drug screening; diagnostic;
KW microorganism detection; potassium channel; fusion protein;
KW multimeric resistance protein 1; MRP1, K1r6.2; mulein.
XX
OS Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX
FH Key
FH Region
FT /note= "Multidrug resistance protein 1 (MRP1)"
FT 1537, 1537
FT /note= "Hexaglycine spacer"
FT 1538, 1927
FT /note= "K1r6.2 mutant protein"
FT M1sc-difference 1907
FT /note= "Wild-type Lys substituted by Ala"
FT M1sc-difference 1908
FT /note= "Wild-type Arg substituted by Ala"
XX
PN US2005063989-A1.
XX
PD 24-MAR-2005.
XX
PF 22-SEP-2003; 2003US-00665283.
XX
PR 22-SEP-2003; 2003US-00665283.
XX
XX (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX
PI Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M,
XX
DR WPI; 2005-252611/26.
XX
XX
PT New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
PS Claim 25; SEQ ID NO 6; 78bp; English.
XX
CC The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multimeric resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel K1r6.2 mutant (KR370AA) protein.
XX
SQ Sequence 1927 AA;
Query Match 100.0%; Score 9734; DB 9; Length 1927;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 121 TFLIQERRKGVQSSGIMLTFMVLVAVCALAIIIRSKIMTALKEDAQVDLPDITFYVYS 180
181 LLLIQLVLSGCFSDRSPLESETIHDNPNCPSSASFLSRITFMWITGLIVRGYRPLEGSD 240
181 LLLIQLVLSGCFSDRSPLESETIHDNPNCPSSASFLSRITFMWITGLIVRGYRPLEGSD 240
241 LMSLNKEDTSEQVVPVLRKWKKECATRQOPVYVYSSQDPAPKSSSKVDANEVEAL 300
241 LMSLNKEDTSEQVVPVLRKWKKECATRQOPVYVYSSQDPAPKSSSKVDANEVEAL 300
301 IVKSPQKEMPSPLEKVLVYKTFGPFLMSFFFKAIHDMFMFSGPOILKLLIKFVNDTAPD 360
301 IVKSPQKEMPSPLEKVLVYKTFGPFLMSFFFKAIHDMFMFSGPOILKLLIKFVNDTAPD 360
361 WQGYFTVLLFVTAQLQTLVLAHQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
361 WQGYFTVLLFVTAQLQTLVLAHQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
421 GEIVNLMVSVDQRFMDLATYINMWSAPLOVIALYLLMNLGSPSVLAGAVVWVLMVFN 480
421 GEIVNLMVSVDQRFMDLATYINMWSAPLOVIALYLLMNLGSPSVLAGAVVWVLMVFN 480
481 AVMAKTKTYQVAHMKSKNRKIKLMEIILNGIKVLKLYAMELAFKDKVLAIROBELKYLK 540
481 AVMAKTKTYQVAHMKSKNRKIKLMEIILNGIKVLKLYAMELAFKDKVLAIROBELKYLK 540
541 KSAVLASAVGTFTWCTPFLVALCTFAVYVITDENNNIIDAQTAFVSLAFNLRPEPLILP 600
541 KSAVLASAVGTFTWCTPFLVALCTFAVYVITDENNNIIDAQTAFVSLAFNLRPEPLILP 600
601 MVISSIVQASVLSKRLRIFLSHEBLEPDSIERRPVKDGGGTSITVANAFTWARSDDPT 660
601 MVISSIVQASVLSKRLRIFLSHEBLEPDSIERRPVKDGGGTSITVANAFTWARSDDPT 660
661 LNGITTSIPRGALVAVVGVGCGKSSLSLALBMDVBEHVAKGSVAVPOQAMQND 720
661 LNGITTSIPRGALVAVVGVGCGKSSLSLALBMDVBEHVAKGSVAVPOQAMQND 720
721 SLRENILFGCQLEPEPYRSVIOACALLPDEILPDSGRTEIGEGVNLGSGQKORVSLAR 780
721 SLRENILFGCQLEPEPYRSVIOACALLPDEILPDSGRTEIGEGVNLGSGQKORVSLAR 780
781 AVYSNADITYLPDPLSANDAHVGHIPENYIGPKMKKNTRIILVTHSMGYLPQVDYIIV 840
781 AVYSNADITYLPDPLSANDAHVGHIPENYIGPKMKKNTRIILVTHSMGYLPQVDYIIV 840
841 MSGKTSSEMGSYOELLARDGAFAPLRTVASTROBDAENGVTGVSQPKKAKOMENGM 900
841 MSGKTSSEMGSYOELLARDGAFAPLRTVASTROBDAENGVTGVSQPKKAKOMENGM 900
901 LVYDSAGKQLOROLSSSSSYSGDISRHNSHTAELOKAAKKEBTWKLMEADKAQOTGVKL 960
901 LVYDSAGKQLOROLSSSSSYSGDISRHNSHTAELOKAAKKEBTWKLMEADKAQOTGVKL 960
961 SVTWDMYKAIGLTISFLSTFLFMCNIVSALASNYMTSLMTDDPIVNGTOGHTYKRLSVYG 1020
961 SVTWDMYKAIGLTISFLSTFLFMCNIVSALASNYMTSLMTDDPIVNGTOGHTYKRLSVYG 1020
1021 ALGISGIAVFGVSMVSTGGILASRCIAVDLHSTILRSFMSFFERTPSGNVNRPSKEL 1080
1021 ALGISGIAVFGVSMVSTGGILASRCIAVDLHSTILRSFMSFFERTPSGNVNRPSKEL 1080
1081 DTVDSMIPEVIKMFMSLFFNVIGACIVILATPAAIIIPPLGLIYFVQRFVASSRQL 1140
1081 DTVDSMIPEVIKMFMSLFFNVIGACIVILATPAAIIIPPLGLIYFVQRFVASSRQL 1140
1141 KRLSEVSRSPVSHFNETLIGSVIRAFEEQEFTHOSDLKVDENQKAYPSIIVARWLA 1200
1141 KRLSEVSRSPVSHFNETLIGSVIRAFEEQEFTHOSDLKVDENQKAYPSIIVARWLA 1200
1201 VRLCEVNCIVLFAALFAVIRSHSLAGVGLSVSYSLQVTVYLNILVMSMETNIVA 1260
1201 VRLCEVNCIVLFAALFAVIRSHSLAGVGLSVSYSLQVTVYLNILVMSMETNIVA 1260

Db 1201 VRLCEVNCIVLFAALPAVVISRHSLSAGLVLSYSISQVTTVYLNWLVMSSEMETNIVA 1260
Qy 1261 VERLKEYSETSEKEAPWQIOETAPSPSWPOVGRVEFRNYCLRYREDLDFVLHINVTINGG 1320
Db 1261 VERLKEYSETSEKEAPWQIOETAPSPSWPOVGRVEFRNYCLRYREDLDFVLHINVTINGG 1320
Qy 1321 EKVGIVGTGAGKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
Db 1321 EKVGIVGTGAGKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
Qy 1381 SSGSRMNUDEPSOVSDEEVMVTSLELAHKDFVSALPDLDEHCAEGGNNLGVGQROLVCL 1440
Db 1381 SSGSRMNUDEPSOVSDEEVMVTSLELAHKDFVSALPDLDEHCAEGGNNLGVGQROLVCL 1440
Qy 1441 ARALLRKTKILVDEATVAVDLETDLDIOSTRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKTKILVDEATVAVDLETDLDIOSTRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Qy 1501 DKGEIOEYGA PSDLIQORGLFYSMAKDAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
Db 1501 DKGEIOEYGA PSDLIQORGLFYSMAKDAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
Qy 1561 PRYTRERARFVSKGNCVNAHKNIROGRFLDOVFTTLVDLKMPHLLIFTMSEFLGSW 1620
Db 1561 PRYTRERARFVSKGNCVNAHKNIROGRFLDOVFTTLVDLKMPHLLIFTMSEFLGSW 1620
Qy 1621 LLFAMVWMLIAFAHGDLPAGEGTNVCVTSIHSFSSAFLEISIEVOVITIGFGRWTEBCP 1680
Db 1621 LLFAMVWMLIAFAHGDLPAGEGTNVCVTSIHSFSSAFLEISIEVOVITIGFGRWTEBCP 1680
Qy 1681 LAIILIVQNVIGLMINAIMGCTFMKTAQAHRAETLIFSKHAVITLRHGLCPMLRAG 1740
Db 1681 LAIILIVQNVIGLMINAIMGCTFMKTAQAHRAETLIFSKHAVITLRHGLCPMLRAG 1740
Qy 1741 DLKRSMTISATIHMOVARKTSPGEVVPPLHNOVDIPMENGNGNCFILVAPLIYHVIDS 1800
Db 1741 DLKRSMTISATIHMOVARKTSPGEVVPPLHNOVDIPMENGNGNCFILVAPLIYHVIDS 1800
Qy 1801 NSPLVDLAPSDLHHQDLEIIVILEGVETTGITTOARTSYLADEILMGORFPIVAEED 1860
Db 1801 NSPLVDLAPSDLHHQDLEIIVILEGVETTGITTOARTSYLADEILMGORFPIVAEED 1860
Qy 1861 GRYSVDYSKFGNTIKVPTPLCTARQLDEDRS 1891
Db 1861 GRYSVDYSKFGNTIKVPTPLCTARQLDEDRS 1891

RESULT 3
ADY86938
ID ADY86938 standard; protein; 1927 AA.
AC ADY86938;
XX 02-JUN-2005 (first entry)
XX Human MRPl-mouse Klr6.2 fusion protein, SEQ ID NO: 1.
XX Ionophore; biosensor; drug screening; diagnostic;
XX microorganism detection; potassium channel; fusion protein;
XX multidrug resistance protein 1; MRPl; Klr6.2.
XX Homo sapiens.
XX Mus musculus.
XX Chimeric.
XX
XX Key Location/Qualifiers
XX Region 1..1531
XX Region /note= "Multidrug resistance protein 1 (MRPl)"
XX Region /note= "Hexaglycine spacer"
XX Region 1538..1927
XX /note= "Klr6.2 protein"

PN US2005063989-A1.
XX 24-MAR-2005.
XX 22-SEP-2003; 2003US-00665283.
XX 22-SEP-2003; 2003US-00665283.
XX (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX Detand R, Garcia E, Prost A, Revilloud J, Vivaudou M,
XX WPI; 2005-252611/26.
XX New ion channel hybrid protein, used as electrical sensor for screening
XX an agonist/antagonist of a membrane protein and for detecting a
XX contaminant/pollutant in a sample.
XX Claim 25; SEQ ID NO 1; 78bp; English.
XX The present invention relates to a hybrid protein consisting essentially
XX of the fusion of a membrane protein with an ion channel which is not
XX naturally coupled to the membrane protein. The hybrid protein is used as
XX an electrical sensor for screening of an agonist/antagonist of a membrane
XX protein, drugs and for detecting a contaminant/pollutant in a sample. The
XX invention is also useful for medical diagnostics and microorganism
XX detection. The present sequence is a fusion protein comprising human
XX multidrug resistance protein 1 (MRPl) and mouse inner rectifier potassium
XX channel Klr6.2 protein.
XX Sequence 1927 AA:
Qy 1 MALRFGSADGSDPLMDMNTVNTNSNDFTYKCFONTVLAWVPCYLLMACPFYFLYLSRH 60
Db 1 MALRFGSADGSDPLMDMNTVNTNSNDFTYKCFONTVLAWVPCYLLMACPFYFLYLSRH 60
Qy 61 DRGYQMTPLNKTKTALGFLIMVQWADLFSPERSRGFLAVFLVSPLLGLITLLA 120
Db 61 DRGYQMTPLNKTKTALGFLIMVQWADLFSPERSRGFLAVFLVSPLLGLITLLA 120
Qy 121 TFLIQLEERKGVSSGIMLTFEVALYCALAILRSKIMTALKEDAQVDLFFRDIFFYVFS 180
Db 121 TFLIQLEERKGVSSGIMLTFEVALYCALAILRSKIMTALKEDAQVDLFFRDIFFYVFS 180
Qy 181 LLLIQVLVSCFSDRSPLESETIHDNPPCESSASFLSRITFWITGLIVRGYRPLSGSD 240
Db 181 LLLIQVLVSCFSDRSPLESETIHDNPPCESSASFLSRITFWITGLIVRGYRPLSGSD 240
Qy 241 LMSLNKEDTSQVVPVLYVKNKKCAKTRKQPVVVVSSKQPAQKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVLYVKNKKCAKTRKQPVVVVSSKQPAQKSSKVDANEVEAL 300
Qy 301 IVKSPOKEWNPSTLEFVLYKTFGPFYFMSFFPKAIDHLMFSGPOLIKLILFVNDTKAPD 360
Db 301 IVKSPOKEWNPSTLEFVLYKTFGPFYFMSFFPKAIDHLMFSGPOLIKLILFVNDTKAPD 360
Qy 361 WQGYFTYVLLFVTACLOTLVLAHQYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFTYVLLFVTACLOTLVLAHQYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
Qy 421 GEIVNLMSVDAQREFMDLATYINMTWSAPLOVITLALYLLMLNLSGSLVAGVAVMTLVVN 480
Db 421 GEIVNLMSVDAQREFMDLATYINMTWSAPLOVITLALYLLMLNLSGSLVAGVAVMTLVVN 480
Qy 481 AVMAKTKTGVVAHKKSKDNRIKILNNEITLNGIKVYLKYAMELAFKDKVLAIROBELKYLK 540
Db 481 AVMAKTKTGVVAHKKSKDNRIKILNNEITLNGIKVYLKYAMELAFKDKVLAIROBELKYLK 540
Qy 541 KSAVLSAVGTFETWCTPPLVALCTPAVVYVITIDENNIIDAQTAFAVSLALFNILRPLNITLP 600

CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 protein containing HA epitope peptide.
XX

Sequence 1947 AA:

Query Match 99.7%; Score 9704; DB 9; Length 1947;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

QY 1 MALMGFSGADSDSLMDMNTWNTSNPDTFCQFQNTLVWVPCFPLMACFPFPLYSRH 60
DB 1 MALMGFSGADSDSLMDMNTWNTSNPDTFCQFQNTLVWVPCFPLMACFPFPLYSRH 60
QY 61 DRGIQMTPLNKTKTALGFLMTICMADLFYSFMERSGIFLAPFLVSPILGITTLLA 120
DB 61 DRGIQMTPLNKTKTALGFLMTICMADLFYSFMERSGIFLAPFLVSPILGITTLLA 120
QY 121 TFLIQLEBRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180
DB 121 TFLIQLEBRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180
QY 181 LLLIQVLSCFSDSPSLFSETHIDNPPCPSSASFLSITFWMTGLIVRGYRQPLEGSD 240
DB 181 LLLIQVLSCFSDSPSLFSETHIDNPPCPSSASFLSITFWMTGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSQOVVPLVANKMKKECAKTRKOPVKVYSSKDPAPKSSKVDANEVVAL 300
DB 241 LMSLNKEDTSQOVVPLVANKMKKECAKTRKOPVKVYSSKDPAPKSSKVDANEVVAL 300
QY 301 IVKSPQKEMNDSLFKVLVYKTFGYPYLMSPFFKAIHDMFMFSGPQILKLIRVNDTKAP 360
DB 301 IVKSPQKEMNDSLFKVLVYKTFGYPYLMSPFFKAIHDMFMFSGPQILKLIRVNDTKAP 360
QY 361 WQGFYTYVLVVTACLOTLVHOYHITCFVSGMRKTAIVIGAVRKALVITNSARKSTV 420
DB 361 WQGFYTYVLVVTACLOTLVHOYHITCFVSGMRKTAIVIGAVRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDQORFMDLATYINMISAPLOYIALYLLMLNGSPVLGAVVMVLMVYN 480
DB 421 GEIVNLMSVDQORFMDLATYINMISAPLOYIALYLLMLNGSPVLGAVVMVLMVYN 480
QY 481 AVMAKTKTYOVAAHKSKDNRIKLNBEILNGIKVLKYAMELAFKQVLAIROBELKVLK 540
DB 481 AVMAKTKTYOVAAHKSKDNRIKLNBEILNGIKVLKYAMELAFKQVLAIROBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAOTAFVSLAFNILRPLNLP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAOTAFVSLAFNILRPLNLP 600
QY 601 MVISIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
DB 601 MVISIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
QY 661 LMGITFSIPBEGALVAVVGQVGGKSSLSALLAEMDKYEGHAIKGSVAYVPOQAMIOND 720
DB 661 LMGITFSIPBEGALVAVVGQVGGKSSLSALLAEMDKYEGHAIKGSVAYVPOQAMIOND 720
QY 721 SLRENILFCQLEBPYRSVITQACALLPDLLETLPESGDRTEIGEKVNLSSGQKQSVSLAR 780
DB 721 SLRENILFCQLEBPYRSVITQACALLPDLLETLPESGDRTEIGEKVNLSSGQKQSVSLAR 780
QY 781 AVYGNADITYLFDPLSAVDAAHVGHIFENNVGPKMKLNKRIILVTHMSYLPQVDVITV 840
DB 781 AVYGNADITYLFDPLSAVDAAHVGHIFENNVGPKMKLNKRIILVTHMSYLPQVDVITV 840
QY 841 MSGGKISMGASYOELLARDGAFAEFLRTYASTEQDQAEENGVTGVSQGEAKQOMENGM 900
DB 841 MSGGKISMGASYOELLARDGAFAEFLRTYASTEQDQAEENGVTGVSQGEAKQOMENGM 900
QY 901 LVTDAGKOLQROUSSSSSYSGDISRHNSTAELOKAEAKKEETWKLMEADKQOTGOYKL 960
DB 901 LVTDAGKOLQROUSSSSSYSGDISRHNSTAELOKAEAKKEETWKLMEADKQOTGOYKL 960

DB 901 LVTDAGKOLQROUSSSSSYSGDISRHNSTAELOKAEAKKEETWKLMEADKQOTGOYKL 960
QY 961 SVYDWYKKAIGLFIISFLIFLFCMCHVSALASNYWLSLWTDDPVINGTOEHTKVLRSYG 1020
DB 961 SVYDWYKKAIGLFIISFLIFLFCMCHVSALASNYWLSLWTDDPVINGTOEHTKVLRSYG 1020
QY 1021 ALGISGIAVFGYSMAVSTIGIILASRCLHYDLHSILSRPMSFPERTPSGULVNRFSKEL 1080
DB 1021 ALGISGIAVFGYSMAVSTIGIILASRCLHYDLHSILSRPMSFPERTPSGULVNRFSKEL 1080
QY 1081 DTVDMSIPEVYKMMGSLFNVIGACIVILATPIAIIIPPLGIYFPVORFYVASSROL 1140
DB 1081 DTVDMSIPEVYKMMGSLFNVIGACIVILATPIAIIIPPLGIYFPVORFYVASSROL 1140
QY 1141 KRLESVSRSPYSHFNSTLLGVSYIRAFEBQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNSTLLGVSYIRAFEBQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
QY 1201 VRLECVGNCIYLPALFAVISRHSLSAGLVLSYSYLOTTYINMLVRRSSEMETNIVA 1260
DB 1201 VRLECVGNCIYLPALFAVISRHSLSAGLVLSYSYLOTTYINMLVRRSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMGOIETAPSSMPOVGRVEFRNYCLRYREDLPVLRIHIVTINGG 1320
DB 1261 VERLKEYSETEKEAPMGOIETAPSSMPOVGRVEFRNYCLRYREDLPVLRIHIVTINGG 1320
QY 1321 EKVGIIVGRTGAGKSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIIVGRTGAGKSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGIIRMLDPPSQYSDEEWTSLELAHKDFVSAIPKLDHECAEGBNLSVGOROYCL 1440
DB 1381 SGIIRMLDPPSQYSDEEWTSLELAHKDFVSAIPKLDHECAEGBNLSVGOROYCL 1440
QY 1441 ARALLRKTILVDEAPAAVDLETDLIQSTIRQFBCVLIATRAHNTIMDTIRVYL 1500
DB 1441 ARALLRKTILVDEAPAAVDLETDLIQSTIRQFBCVLIATRAHNTIMDTIRVYL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLVGGGGGGLSRKGIIPBEVYLTRIAEDPAE 1560
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLVGGGGGGLSRKGIIPBEVYLTRIAEDPAE 1560
QY 1561 PRYTRERRARFVSKKGCNVAAHNIIBQGRFLQDVFTTVLDLKWPHLLIFTMSFLCSW 1620
DB 1561 PRYTRERRARFVSKKGCNVAAHNIIBQGRFLQDVFTTVLDLKWPHLLIFTMSFLCSW 1620
QY 1621 LLEFAMWMLTAFAG-----DLAP-----GBGVNVCVTSIHSSSAFLF 1660
DB 1621 LLEFAMWMLTAFAGDLVAYMEKGITDLAPYPVDPVAGGNTVPCVTSIHSSSAFLF 1660
QY 1661 SIEVQVTIGFGRAWTEECPLAILILIVONIVGMINAIMGCI FMTQAQHRBAETLIF 1720
DB 1661 SIEVQVTIGFGRAWTEECPLAILILIVONIVGMINAIMGCI FMTQAQHRBAETLIF 1720
QY 1721 SKHAVITLRHGRCLCFMLRVGDLRKSMIISATIHQVVRKTTSPGEVVPPLHVDI PMENG 1780
DB 1721 SKHAVITLRHGRCLCFMLRVGDLRKSMIISATIHQVVRKTTSPGEVVPPLHVDI PMENG 1780
QY 1781 VGGNGIFLVAPLLIYHYIDSNSPIYDLAPSLHHHODELIIIVIEGVETGITTOARTS 1840
DB 1781 VGGNGIFLVAPLLIYHYIDSNSPIYDLAPSLHHHODELIIIVIEGVETGITTOARTS 1840
QY 1841 YLADEILMGQRFPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQOLDEDRS 1891
DB 1841 YLADEILMGQRFPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQOLDEDRS 1891
RESULT 5
AAM57486
ID AAM57486 standard; protein; 1531 AA.
XX
AC AAM57486;
XX

DT 14-AUG-1998 (first entry)
XX Human MRP variant lcp9a (lei/9p9a).
XX Multidrug resistance-associated protein; MRP; tumour; human; variant;
KM multidrug resistance; MDR; leismania P-glycoprotein; lcp9a; lei/9p9a.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 685 /label= L685S
FT /note= "wild-type Leu is replaced by Ser"
FT Misc-difference 1282 /label= R1282A
FT /note= "wild-type Arg is replaced by Ala"
XX
PN US5766880-A.
XX
PD 16-JUN-1998.
XX
PF 05-JUN-1995; 95US-00463092.
XX
PR 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
PA (TOOH) UNIV QUEBENS KINGSTON.
XX
PI Cole SP, Deeley RG;
XX
DR WPI: 1998-361687/31.
DR N-PSDB; AAV31498.
XX
XX DNA encoding protein associated with multi-drug resistance - useful for
PT as probe for identifying multi-drug resistant tumour cells.
XX
XX
XX Claim 1; Col 67-78; 82pp; English.
XX
XX This represents a variant of the human multidrug resistance-associated
CC protein (MRP). This natural variant is a leismania P-glycoprotein related
CC molecule lcp9a (lei/9p9a). The human and murine MRP nucleic acid
CC molecules can be used as probes for identifying multidrug resistant
CC tumour cells by hybridisation to mRNA from tumour cells. The antisense
CC nucleic acid can be used to reverse multidrug resistance (MDR). A
CC recombinant expression vector containing the MRP nucleic acid molecules
CC operatively linked to at least one regulatory sequence can be used to
CC transform a host cell to produce a recombinant MDR-associated protein
XX
SQ Sequence 1531 AA;
Query Match 80.7%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 LMSLNKEDTSEQVAVLVKWKKECAKTRKQPVVAVVSSKDPAQPKSSSKVDANEVEAL 300
QY 301 IVSPQKKNPSSLEKVLKTFPGPYFLMSFFPKAHDLMFSGPOILKLLIKFVNDTAAPD 360
DB 301 IVSPQKKNPSSLEKVLKTFPGPYFLMSFFPKAHDLMFSGPOILKLLIKFVNDTAAPD 360
QY 361 WQGFYFVLLPVTAQCTVLVHOVPHICFVSGMKIKTAIVGAVRKALVITNSARKSTV 420
DB 361 WQGFYFVLLPVTAQCTVLVHOVPHICFVSGMKIKTAIVGAVRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRMEDATYINMWSAPLOVIALYLLMNLGSPVLAVAVMLAMYPVN 480
DB 421 GEIVNLMSVDAQRMEDATYINMWSAPLOVIALYLLMNLGSPVLAVAVMLAMYPVN 480
QY 481 AVNAAKTTYQVAMKSKDNRIKLMNEILNGIKVLKYAMELAKDVKVLAIRBELKVLK 540
DB 481 AVNAAKTTYQVAMKSKDNRIKLMNEILNGIKVLKYAMELAKDVKVLAIRBELKVLK 540
QY 541 KSAVLAVGTFPTWCPELVALCTFAVYVITDENNILDQAFAVSLAFNLRPEPLILP 600
DB 541 KSAVLAVGTFPTWCPELVALCTFAVYVITDENNILDQAFAVSLAFNLRPEPLILP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELBPSIERRPVKDGGTNSITVBNATFTMARSDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELBPSIERRPVKDGGTNSITVBNATFTMARSDPT 660
QY 661 LNCITFSPGALVAVVGVCCKSSLLSALLAEMDRVGHVAIKGSVAAYVPOQAWI QND 720
DB 661 LNCITFSPGALVAVVGVCCKSSLLSALLAEMDRVGHVAIKGSVAAYVPOQAWI QND 720
QY 721 SLRENILFGQLEBPYRSVIOACALLPDLIELPSSGRTIEGKGVNLSSGQKRVSLAR 780
DB 721 SLRENILFGQLEBPYRSVIOACALLPDLIELPSSGRTIEGKGVNLSSGQKRVSLAR 780
QY 781 AVVSNADITYLPDDPLSAVDAAVGHIPENYIGPKMKAKNTRILVITSMSTLPQVDYIIV 840
DB 781 AVVSNADITYLPDDPLSAVDAAVGHIPENYIGPKMKAKNTRILVITSMSTLPQVDYIIV 840
QY 841 MSGGKISEMWSYOELARDGAPAEFLRTVASTEOBDAEENGVTGVSPPCKEAKOMENGM 900
DB 841 MSGGKISEMWSYOELARDGAPAEFLRTVASTEOBDAEENGVTGVSPPCKEAKOMENGM 900
QY 901 LVYDSACKQIQORQLSSSSSSISGDISRRHNSTAEIQKAEKKEETWKLMEADKAQTGVKL 960
DB 901 LVYDSACKQIQORQLSSSSSSISGDISRRHNSTAEIQKAEKKEETWKLMEADKAQTGVKL 960
QY 961 SVYWDYKAKIGLFIISFLSIFLFCMCHVSAIASNYMISLMTDDPIVNGTOSEHTKVRISVYG 1020
DB 961 SVYWDYKAKIGLFIISFLSIFLFCMCHVSAIASNYMISLMTDDPIVNGTOSEHTKVRISVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHLSILRSPMSFEERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHLSILRSPMSFEERTPSGNLVNRFSEKL 1080
QY 1081 DTWDSMIPETIYKMFNGSLFNVIACIYIILATPIAIIIPPLGLIYFVORFVYASRQL 1140
DB 1081 DTWDSMIPETIYKMFNGSLFNVIACIYIILATPIAIIIPPLGLIYFVORFVYASRQL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEBQERFIHOSDLKVDENQAKAYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEBQERFIHOSDLKVDENQAKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCTVLPFAALFAVISRSLSAGLVGSVSYSLQVTTYIANMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVGNCTVLPFAALFAVISRSLSAGLVGSVSYSLQVTTYIANMLVRMSSEMETNIVA 1260
QY 1261 VERLXYSERTEKAPQIOETAPSSWPQGRVREPNYICLRREDDDFVLRHNVITNGG 1320
DB 1261 VERLXYSERTEKAPQIOETAPSSWPQGRVREPNYICLRREDDDFVLRHNVITNGG 1320
QY 1321 EKVGIVGRTGAGKSLTLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380

Db 1321 EKVGIVGRTGAGKSLTLGLFRINESAGEIITIGINIAKIGLHDLRPRKTIIPDPVLF 1380
QY 1381 SGLSRMNLDPFSQXSDEEVTWTSLELAHLKDPVSALPDLDHECARGENLSVGOROLVCL 1440
Db 1381 SGLSRMNLDPFSQXSDEEVTWTSLELAHLKDPVSALPDLDHECARGENLSVGOROLVCL 1440
QY 1441 ARALLRKTKIIVLDEATPAVDLETFDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKTKIIVLDEATPAVDLETFDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOYVGAAPSDLLQORGLFYMAKXAGLV 1531
Db 1501 DKGEIOYVGAAPSDLLQORGLFYMAKXAGLV 1531
RESULT 6
ID AAW74471 standard; protein, 1531 AA.
XX AAW74471;
DT 18-MAY-1999 (first entry)
XX Human multidrug resistance-associated protein variant.
XX Multidrug resistance-associated protein; MDR; human; diagnosis;
KM MDR tumour cell identification; cancer therapy.
XX Homo sapiens.
XX US5882875-A.
PN 16-MAR-1999.
PD 16-MAR-1999.
XX 05-JUN-1995; 95US-00462109.
XX 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX (TOOH) UNIV QUEBENS KINGSTON.
PA Cole SPC, Deeley RG;
PI WPI: 1999-214061/18.
DR N-PSDB; AAX21977.
XX Identifying a multidrug resistant tumor cell by contacting the cell with
PT an antibody/antigen-binding fragment - which binds to an expressed
PT protein encoded by multidrug resistance-associated protein (MRP) nucleic
PT acid.
XX Claim 3; Col 69-80; 80pp; English.
PS This sequence is the human multidrug resistance-associated (MDR) protein.
XX The invention relates to a method for identifying a multidrug resistant
CC (MDR) tumour cell. Compositions and methods utilising the MDR proteins
CC can be used to treat patients with tumours displaying multidrug
CC resistance, particularly those displaying resistance to anthracyclines,
CC epipodophylotoxins, vinca alkaloids, and hydrophobic drugs. The methods
CC for inhibiting/killing a MDR tumour cell can be useful for treating
CC breast cancer, leukemias, fibrosarcomas, cervical cancer, gliomas
CC thymomas, neuroblastomas and lung cancer. The MDR DNA sequences when
CC labeled are useful as molecular probes for diagnosing multidrug
CC resistance of a tumour (using cells from a tumour biopsy) and for
CC designing ribozymes which are capable of cleaving a single-stranded
CC nucleic acid encoding a protein having MRP activity. Recombinant
CC expression vectors containing human MDR coding sequences can be
CC transfected into a drug sensitive cell line to produce a protein in the
CC cell which confers MDR, protecting non-resistant non-tumour cells from
CC the effects of chemotherapeutics has major clinical importance. Cells
CC transformed with the MDR coding sequences are useful for testing

CC potential therapeutic agents for their effectiveness against MDR cells
CC and for identifying chemosensitizers of a therapeutic agent
XX
SQ Sequence 1531 AA;
Query Match 80.7%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSAGSDPDLMDMNTWNTSNPDFFKCFONTLVWVPCFYLMACPFYFLYLSRH 60
Db 1 MALRGFCSAGSDPDLMDMNTWNTSNPDFFKCFONTLVWVPCFYLMACPFYFLYLSRH 60
QY DRGIQMTPLNKTKTALGFLIMIVCMADLFYSFWEBSRGIFLAVPLVSPITLIGITTLA 120
Db DRGIQMTPLNKTKTALGFLIMIVCMADLFYSFWEBSRGIFLAVPLVSPITLIGITTLA 120
QY 121 TFLIQLEBRKGVSSGIMLTFWLVALYCALAILRSKIMTALKEDAOVDLPFDITFYVYS 180
Db 121 TFLIQLEBRKGVSSGIMLTFWLVALYCALAILRSKIMTALKEDAOVDLPFDITFYVYS 180
QY 181 LLLIQLVLSGSDPSPLFSETIHDNPPCPSSASFSLRITPFMTITGLIVRGYRQPLSGSD 240
Db 181 LLLIQLVLSGSDPSPLFSETIHDNPPCPSSASFSLRITPFMTITGLIVRGYRQPLSGSD 240
QY 241 LMSLNKEDTSEQVVPVLYKWKKECAKTRKQPVKVVYSSKDPAPKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLYKWKKECAKTRKQPVKVVYSSKDPAPKSSKYDANEVEAL 300
QY 301 IVKSPQKBNPDLFKVLYKTFGPFLMSFPFKAIHDLMSRGPOTLKLLIKFVNDTPAD 360
Db 301 IVKSPQKBNPDLFKVLYKTFGPFLMSFPFKAIHDLMSRGPOTLKLLIKFVNDTPAD 360
QY 361 WQGFYTVLLEFVTAQCTQIVLHOYFHI CFVSGMS IKTAVGAVYRKALVITNSARKSTV 420
Db 361 WQGFYTVLLEFVTAQCTQIVLHOYFHI CFVSGMS IKTAVGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMNLGSPVLAVGAVWVLMVNV 480
Db 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMNLGSPVLAVGAVWVLMVNV 480
QY 481 AVMAMKTKTYQVAMHKKDNRIKLMNEILINGIKVLKYAMELAKDKVLAIROBELKVLK 540
Db 481 AVMAMKTKTYQVAMHKKDNRIKLMNEILINGIKVLKYAMELAKDKVLAIROBELKVLK 540
QY 541 KSAVLSAVGFTWCTPEFVALCTFAVYVITDENNNIIDAQTAFFSLALFNLRPLNLTLP 600
Db 541 KSAVLSAVGFTWCTPEFVALCTFAVYVITDENNNIIDAQTAFFSLALFNLRPLNLTLP 600
QY 601 MVISIVQASVSLKRLHIFLSHEELPDSIERRPVKDGGTNSITVNAATPTWARSDEPT 660
Db 601 MVISIVQASVSLKRLHIFLSHEELPDSIERRPVKDGGTNSITVNAATPTWARSDEPT 660
QY 661 LNGITPSIPBEGALVAVVGVQVCGKSSLLSALLAMDEKVEGHVAKGSAAVYPOQAMIOND 720
Db 661 LNGITPSIPBEGALVAVVGVQVCGKSSLLSALLAMDEKVEGHVAKGSAAVYPOQAMIOND 720
QY 721 SLRNNILFGCOLLEPPYRSVYQAQALLPDLEILSGBRTEIGEKNVLSGGQKRVSLAR 780
Db 721 SLRNNILFGCOLLEPPYRSVYQAQALLPDLEILSGBRTEIGEKNVLSGGQKRVSLAR 780
QY 781 AAVSNADIVLPDDPLSAVDAAVGHIFENYVGPKGMLKNKRIIVTSHMSYLPQVDVIIV 840
Db 781 AAVSNADIVLPDDPLSAVDAAVGHIFENYVGPKGMLKNKRIIVTSHMSYLPQVDVIIV 840
QY 841 MSGKISMSGSYQILARDAFAFLRTYASTEOBDAEENGVTGVSQPGKBAKOMENG 900
Db 841 MSGKISMSGSYQILARDAFAFLRTYASTEOBDAEENGVTGVSQPGKBAKOMENG 900
QY 901 LVTPSAGKOLOROUSSSSSYGDISRHNTGTAELQXAKKEEPMWKLMEADKAOYGVL 960
Db 901 LVTPSAGKOLOROUSSSSSYGDISRHNTGTAELQXAKKEEPMWKLMEADKAOYGVL 960

QY 961 SVYDYMKAIGLFLISFLISFLMCMNHVSALASNYWLSMTDDPIVNGTOEHTKYRLSVYG 1020
|||
DB 961 SVYDYMKAIGLFLISFLISFLMCMNHVSALASNYWLSMTDDPIVNGTOEHTKYRLSVYG 1020
QY 1021 ALGISOGIANGVGMANVISIGIILASRCLHVDLHSLILSPMSFFERTPSGNLVNRFSEKL 1080
|||
DB 1021 ALGISOGIANGVGMANVISIGIILASRCLHVDLHSLILSPMSFFERTPSGNLVNRFSEKL 1080
QY 1081 DTYSWMIPEVTKMGMGSLFNIVIGACIVLTAATPIAIIIPPLGLIYFVQGFYVASSKQL 1140
|||
DB 1081 DTYSWMIPEVTKMGMGSLFNIVIGACIVLTAATPIAIIIPPLGLIYFVQGFYVASSKQL 1140
QY 1141 KRLESVSRSPYSHFNELLGVSVIRAFEEQERFIHOSDLKYDENOKAYYSIVANRWLA 1200
|||
DB 1141 KRLESVSRSPYSHFNELLGVSVIRAFEEQERFIHOSDLKYDENOKAYYSIVANRWLA 1200
QY 1201 VRLFCVGNCLVLPALPAVISRHSLSAGLVGLSVSYSLQVTTYINWLVMSSEMETNIVA 1260
|||
DB 1201 VRLFCVGNCLVLPALPAVISRHSLSAGLVGLSVSYSLQVTTYINWLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIQETAPSSWPQVGRVFRNYCLAYREDLPVLAHINVTINGG 1320
|||
DB 1261 VERLKEYSETEKAPWQIQETAPSSWPQVGRVFRNYCLAYREDLPVLAHINVTINGG 1320
QY 1321 EKVGIVERTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRPKTIIPQDPVLF 1380
|||
DB 1321 EKVGIVERTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRPKTIIPQDPVLF 1380
QY 1381 SGIIRKMLDPPSQYSDEEWTSLSLAHLKDFVSALPDKLDHECAEGENLSVGQROLVCL 1440
|||
DB 1381 SGIIRKMLDPPSQYSDEEWTSLSLAHLKDFVSALPDKLDHECAEGENLSVGQROLVCL 1440
QY 1441 ARALLRTKTLIVDEATPAVDLETDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRYIVL 1500
|||
DB 1441 ARALLRTKTLIVDEATPAVDLETDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRYIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGLV 1531
|||
DB 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGLV 1531

RESULT 7
AAM9894
ID AAM9894 standard; protein: 1531 AA.
XX
AC AAM9894;
XX
DT 10-JUN-1999 (first entry)
XX
DE Human multidrug resistance-associated protein natural variant.
XX
KW Human; multidrug resistance-associated protein; MRP; cytotoxic drug;
KM cancer; chemotherapy.
XX
OS Homo sapiens.
XX
PN USS891724-A.
XX
PD 06-APR-1999.
XX
PF 05-JUN-1995; 95US-00460907.
XX
PR 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
PA (TOOH) UNIV QUEBENS KINGSTON.
XX
PI Cole SPC, Deeley RG;
XX
XX WPI; 1999-253868/21.
DR N-PSDB; AAX19818.

XX
PT Protecting mammalian cells against cytotoxic drugs.
XX
PS Claim 2; Col 79-86; 82pp; English.
XX
CC The present sequence represents a human multidrug resistance-associated
CC protein (MRP). The present invention also describes a method for
CC protecting a mammalian cell against the cytotoxicity of anthracyclines,
CC epipodophyllotoxins and vinca alkaloids (A) by introducing into it a
CC nucleic acid (I) that hybridizes under stringent conditions to a nucleic
CC acid (II) that encodes an MRP protein (III). Introduction of (I) protects
CC cells against cytotoxic effects of (A), particularly to protect normal
CC cells against (A) being used for treatment of cancers. Cells transformed
CC with (I) can be used to screen for agents that affect multidrug
CC resistance or are directly toxic to multidrug resistant cells, i.e.
CC potential therapeutics for multidrug-resistant cancers. Confering
CC resistance to normal cells should allow an increase in the dose of (A)
CC that can be administered safely

SO Sequence 1531 AA;

Query Match 80.7%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSPDLMNMTNTNSNDFTKCFQNTLVVWVPCYILMACPPFYLYSRH 60
1 MALRGFCSADGSPDLMNMTNTNSNDFTKCFQNTLVVWVPCYILMACPPFYLYSRH 60
DB 1 MALRGFCSADGSPDLMNMTNTNSNDFTKCFQNTLVVWVPCYILMACPPFYLYSRH 60
QY 61 DRGYIQMTPLNKTALGFLMTIVCMADLFYSFMRSGIFLAPFVSPFLIGITTLA 120
61 DRGYIQMTPLNKTALGFLMTIVCMADLFYSFMRSGIFLAPFVSPFLIGITTLA 120
DB 61 DRGYIQMTPLNKTALGFLMTIVCMADLFYSFMRSGIFLAPFVSPFLIGITTLA 120
QY 121 TFLIQERRKGVSSGIMLTFWLVALYCALAIIIRSKIMTAKEDAQVDFRDIIFYVYS 180
121 TFLIQERRKGVSSGIMLTFWLVALYCALAIIIRSKIMTAKEDAQVDFRDIIFYVYS 180
DB 121 TFLIQERRKGVSSGIMLTFWLVALYCALAIIIRSKIMTAKEDAQVDFRDIIFYVYS 180
QY 181 LLLIQVLSGFSRSPFLFSETHIDPNPCSSASFLSRITFWMTGLIYNGROPLEGS 240
181 LLLIQVLSGFSRSPFLFSETHIDPNPCSSASFLSRITFWMTGLIYNGROPLEGS 240
DB 181 LLLIQVLSGFSRSPFLFSETHIDPNPCSSASFLSRITFWMTGLIYNGROPLEGS 240
QY 241 LMSLNKEDTSEOVVPLVVKMKKECAKTRQOPKVVYSSKDPAOPKSSKYDANEVEAL 300
241 LMSLNKEDTSEOVVPLVVKMKKECAKTRQOPKVVYSSKDPAOPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEOVVPLVVKMKKECAKTRQOPKVVYSSKDPAOPKSSKYDANEVEAL 300
QY 301 IVKSPQKWNPSLFFKVLKTFGPFYFLMSFFPKAIDHLMFSGPQILKLLIKFVNDTYAP 360
301 IVKSPQKWNPSLFFKVLKTFGPFYFLMSFFPKAIDHLMFSGPQILKLLIKFVNDTYAP 360
DB 301 IVKSPQKWNPSLFFKVLKTFGPFYFLMSFFPKAIDHLMFSGPQILKLLIKFVNDTYAP 360
QY 361 WQGYFTYVLLFVPAACIOTLVLAHOYFHCFSGMRKITAVIGAVYRKALVTTNSARKSSTV 420
361 WQGYFTYVLLFVPAACIOTLVLAHOYFHCFSGMRKITAVIGAVYRKALVTTNSARKSSTV 420
DB 361 WQGYFTYVLLFVPAACIOTLVLAHOYFHCFSGMRKITAVIGAVYRKALVTTNSARKSSTV 420
QY 421 GEIVNLSVDAORFMDATYINNIWSAPLOVITLALVLYMLNLPVSAGAVAWVLWVNVN 480
421 GEIVNLSVDAORFMDATYINNIWSAPLOVITLALVLYMLNLPVSAGAVAWVLWVNVN 480
DB 421 GEIVNLSVDAORFMDATYINNIWSAPLOVITLALVLYMLNLPVSAGAVAWVLWVNVN 480
QY 481 AVNANKTKTYQVAMKSKNRIKLMEIILNGIKVLYKLAWELAFKQVLAIRQELKVLK 540
481 AVNANKTKTYQVAMKSKNRIKLMEIILNGIKVLYKLAWELAFKQVLAIRQELKVLK 540
DB 481 AVNANKTKTYQVAMKSKNRIKLMEIILNGIKVLYKLAWELAFKQVLAIRQELKVLK 540
QY 541 KSAVLSAVGTFVTCPPFLVALCTFAVYVITIDENNIDAOATAVSLAPFIIAPLNTLP 600
541 KSAVLSAVGTFVTCPPFLVALCTFAVYVITIDENNIDAOATAVSLAPFIIAPLNTLP 600
DB 541 KSAVLSAVGTFVTCPPFLVALCTFAVYVITIDENNIDAOATAVSLAPFIIAPLNTLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHBELEPDSIERPVDGGGTSNITVRNAFTTARSDEPT 660
601 MVISSIVQASVSLKRLRIFLSHBELEPDSIERPVDGGGTSNITVRNAFTTARSDEPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHBELEPDSIERPVDGGGTSNITVRNAFTTARSDEPT 660
QY 661 LMGITFSIPRGALVAVVGQVCGKSSLSALLAEKMDVEGHVAIKGSVAVVPOOAMIOND 720
661 LMGITFSIPRGALVAVVGQVCGKSSLSALLAEKMDVEGHVAIKGSVAVVPOOAMIOND 720
DB 661 LMGITFSIPRGALVAVVGQVCGKSSLSALLAEKMDVEGHVAIKGSVAVVPOOAMIOND 720

```

Qy 721 SLRENIIFGCOLLEPYRVSIVQACALLPDLIELPSGDRTEIGEGKVNISGGQKORVSLAR 780
Db 721 SLRENIIFGCOLLEPYRVSIVQACALLPDLIELPSGDRTEIGEGKVNISGGQKORVSLAR 780
Qy 781 AAVENADIYLPDDLSAVDAHVGHIPENVIQPGMKLNKTRILLVTHSMSTLPQVDVIIV 840
Db 781 AAVENADIYLPDDLSAVDAHVGHIPENVIQPGMKLNKTRILLVTHSMSTLPQVDVIIV 840
Qy 841 MSGGKISEMGSYQELLARDAFAEFLRTYASTEOQDAEENGVTGVSPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYQELLARDAFAEFLRTYASTEOQDAEENGVTGVSPGKEAKOMENGM 900
Qy 901 LVTDASAGKOLQROUSSSSSSYSGDISRHNSTAELOKAEKKEETWKLMEADKAOGTQYKL 960
Db 901 LVTDASAGKOLQROUSSSSSSYSGDISRHNSTAELOKAEKKEETWKLMEADKAOGTQYKL 960
Qy 961 SVYWDYMKATIGLFTISFLSIFLMCHVSAALSNWLSMTDPPVNGQOETHKRLSYVG 1020
Db 961 SVYWDYMKATIGLFTISFLSIFLMCHVSAALSNWLSMTDPPVNGQOETHKRLSYVG 1020
Qy 1021 ALGISOGIAVPGYMAVSIIGILASRCLHVDLHSHILSPMSFFERTPSGNLVNRFKEL 1080
Db 1021 ALGISOGIAVPGYMAVSIIGILASRCLHVDLHSHILSPMSFFERTPSGNLVNRFKEL 1080
Qy 1081 DTVDSMIPEVIKMGSLFNVIAGCTIVILATPIAIIIPPLGLIFFVQRPVYASSRQL 1140
Db 1081 DTVDSMIPEVIKMGSLFNVIAGCTIVILATPIAIIIPPLGLIFFVQRPVYASSRQL 1140
Qy 1141 KRLSVSRPVYSHPNFTLLGVSVIRAFEEQERFHQSDLVKDEQKAYPSIVANRWLA 1200
Db 1141 KRLSVSRPVYSHPNFTLLGVSVIRAFEEQERFHQSDLVKDEQKAYPSIVANRWLA 1200
Qy 1201 VRLCEVGCIVLFAVIRSRHSAGLVGSYSIQVTTYLNMLVRMSEMETNIVA 1260
Db 1201 VRLCEVGCIVLFAVIRSRHSAGLVGSYSIQVTTYLNMLVRMSEMETNIVA 1260
Qy 1261 VERLKEYSETEKAPWQIQETAPPSWPQVGRVEFRNYCLNRREDLDFVLHINVTINGG 1320
Db 1261 VERLKEYSETEKAPWQIQETAPPSWPQVGRVEFRNYCLNRREDLDFVLHINVTINGG 1320
Qy 1321 EKVGIQVGTGAGKSLTGLFRINSEAGEIITIOINAKIGHLRKRTIIPDDPLF 1380
Db 1321 EKVGIQVGTGAGKSLTGLFRINSEAGEIITIOINAKIGHLRKRTIIPDDPLF 1380
Qy 1381 SGLSRMNLDPSSQXSDEEVMTSLLELAHLKDFVSALPDKLDHECAGGENTL SVGORQLVCL 1440
Db 1381 SGLSRMNLDPSSQXSDEEVMTSLLELAHLKDFVSALPDKLDHECAGGENTL SVGORQLVCL 1440
Qy 1441 ARALLRKTKIIVLDEATAVADLETDLLIQSTIRTOFEDCTVLTJAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTKIIVLDEATAVADLETDLLIQSTIRTOFEDCTVLTJAHRLNTIMDYTRIVL 1500
Qy 1501 DKGEIQEYGAPSDLLQORGLFYMAKXKAGLV 1531
Db 1501 DKGEIQEYGAPSDLLQORGLFYMAKXKAGLV 1531

RESULT 8
AAY55799
ID AAY55799 standard; protein; 1531 AA.
AC AAY55799;
AA 28-FEB-2000 (first entry)
DE Human multidrug resistance-associated protein (MRP) variant.
DB Human multidrug resistance-associated protein; MRP; human;
KW chemosensitizer; multidrug resistance-associated protein; MRP; human;
KW therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;
KW cancer; variant.
XX Homo sapiens.
OS
XX

```

```

FH Key Location/Qualifiers
FT Misc-difference 685
FT /label= L685S
FT /note= "wild-type Leu is replaced by Ser"
FT Misc-difference 1282
FT /label= R1282A
FT /note= "wild-type Arg is replaced by Ala"
PN US6001563-A.
XX
XX 14-DEC-1999.
PD
XX
XX 05-JUN-1995; 95US-00463179.
PE
XX
XX 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
XX (TOOH ) UNIV QUEBENS KINGSTON.
PA
XX
XX Cole SP, Deeley RG;
PI WPI; 2000-061877/05.
XX
XX N-PSDB; AAZ39556.
DR
XX Identification of chemosensitizers useful for treating cancer, using
PT nucleic acid encoding multidrug resistance-associated protein.
XX
XX Claim 3; Col 69-80; 77pp; English.
PS
XX
XX The invention provides a method for identifying a substance which is a
CC chemosensitizer that comprises, contacting a cell transfected with
CC nucleic acid encoding multidrug resistance-associated protein (MRP) with
CC a therapeutic agent in vitro. The method is useful for identifying
CC chemosensitizers which may then be used to treat cancer (especially lung
CC cancer). The method allows the identification of chemosensitizers which
CC do not reverse P-glycoprotein-mediated multidrug resistance. The present
CC sequence represents a human MRP variant
XX
SQ Sequence 1531 AA:
Query Match 80.7%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRGFSGADSGDPLMDMNNVTWNTSNPDKCFQNTVLVWVPCFYLMACFPFYFLYLSRH 60
Db 1 MALRGFSGADSGDPLMDMNNVTWNTSNPDKCFQNTVLVWVPCFYLMACFPFYFLYLSRH 60
Qy 61 DRGYIQMTPLNKTKTALGFLIMVCMADLFYSFWERSRGFLIAPVFLVSPTLIGITLLA 120
Db 61 DRGYIQMTPLNKTKTALGFLIMVCMADLFYSFWERSRGFLIAPVFLVSPTLIGITLLA 120
Qy 121 TFLIQLERRKGVSSGIMLTFWLVALVCAALILSKIMTALKEDAQVDLFFRDIIFYVYS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVALVCAALILSKIMTALKEDAQVDLFFRDIIFYVYS 180
Qy 181 LLLIQVLVSCFSDSPFLFSEITHDPNCPRESSAFLSRITFWMTTGLIVRYRQPLBESD 240
Db 181 LLLIQVLVSCFSDSPFLFSEITHDPNCPRESSAFLSRITFWMTTGLIVRYRQPLBESD 240
Qy 241 LMSLINKEDTSQVVPVLYKMKKECAKTRKQPVVYVSSKDPAPKESKSDANBEVVAL 300
Db 241 LMSLINKEDTSQVVPVLYKMKKECAKTRKQPVVYVSSKDPAPKESKSDANBEVVAL 300
Qy 301 IVKSPQKEMNSLRFKVLVYKTFGPFYFLMSFFPKAIDLMFSGPOLKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNSLRFKVLVYKTFGPFYFLMSFFPKAIDLMFSGPOLKLLIKFVNDTKAPD 360
Qy 361 WQGYFYTVLFLVTAQCLQTLVLYHQYFHICFVSGMRKIKTAIVGAVVRKALVITNSARKSSTV 420
Db 361 WQGYFYTVLFLVTAQCLQTLVLYHQYFHICFVSGMRKIKTAIVGAVVRKALVITNSARKSSTV 420

```

```

QY 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVIALYILMLNLSGVSILAGVAVMYLAMPVN 480
DB 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVIALYILMLNLSGVSILAGVAVMYLAMPVN 480
QY 481 AVMAKTKTYQVAMHKSNDRIKLMNEILNGIKLVYAMELAFKDVLAIROBELKYLK 540
DB 481 AVMAKTKTYQVAMHKSNDRIKLMNEILNGIKLVYAMELAFKDVLAIROBELKYLK 540
QY 541 KSAVLASVGTFTWCTPPLVALCTPAYVYTIIDENNILDAQAFVSLAFNLRPLNLP 600
DB 541 KSAVLASVGTFTWCTPPLVALCTPAYVYTIIDENNILDAQAFVSLAFNLRPLNLP 600
QY 601 MVISIVQASVSLKRLRIFLSHEELBPDSIERRPVKOGGNGNSTIVNATATMARSPPT 660
DB 601 MVISIVQASVSLKRLRIFLSHEELBPDSIERRPVKOGGNGNSTIVNATATMARSPPT 660
QY 661 LMGITFSPREGALVAVVGVCGKSSLSALLAEMDKVGHVAKGSVAVYPOQAWIOND 720
DB 661 LMGITFSPREGALVAVVGVCGKSSLSALLAEMDKVGHVAKGSVAVYPOQAWIOND 720
QY 721 SLRENILFGCOLLEPPYRSVIOACALLPDLEILPSGDRTEIGEKGVNLSSGQKQVSLAR 780
DB 721 SLRENILFGCOLLEPPYRSVIOACALLPDLEILPSGDRTEIGEKGVNLSSGQKQVSLAR 780
QY 781 AVYSNADIVLPDDLSAVDAHVGHIFENVIGPKMLKNTKRIIVTHSMSTLPOVDVYIV 840
DB 781 AVYSNADIVLPDDLSAVDAHVGHIFENVIGPKMLKNTKRIIVTHSMSTLPOVDVYIV 840
QY 841 MSGGKISEMGSYOELIARDGAFAPFLTYASTEOBDAEENGVTGSGPGKEAQMENG 900
DB 841 MSGGKISEMGSYOELIARDGAFAPFLTYASTEOBDAEENGVTGSGPGKEAQMENG 900
QY 901 LVTDSAGKQLORLSSSSSYSGDISRHHNSTAELOKAEKKEFTWKLMEADKAQTGVKL 960
DB 901 LVTDSAGKQLORLSSSSSYSGDISRHHNSTAELOKAEKKEFTWKLMEADKAQTGVKL 960
QY 961 SVVWDYKAKIGLFISESTIFLPMCNHVSALASNTWLSMTDDPIVNGTOETKRLSVYG 1020
DB 961 SVVWDYKAKIGLFISESTIFLPMCNHVSALASNTWLSMTDDPIVNGTOETKRLSVYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCIHDVLSIIRSPMSFERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCIHDVLSIIRSPMSFERTPSGNLVNRFSEKL 1080
QY 1081 DTVDMSIPEVYKMFMSGLFNVIACIYILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
DB 1081 DTVDMSIPEVYKMFMSGLFNVIACIYILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNFTLLGVSVIRAFEEBERFIHOSDKVDNOKAYPSIYANWMLA 1200
DB 1141 KRLSVSRSPVYSHFNFTLLGVSVIRAFEEBERFIHOSDKVDNOKAYPSIYANWMLA 1200
QY 1201 VRLCEVCNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETIVA 1260
DB 1201 VRLCEVCNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVIGIVRTGAGKSSLLGLFRINESAGEIIIDGINIAIKGLHDLRFKTIIPQDVPVL 1380
DB 1321 EKVIGIVRTGAGKSSLLGLFRINESAGEIIIDGINIAIKGLHDLRFKTIIPQDVPVL 1380
QY 1381 SSGSLRMLNLDPPSOYSDDEWVTSLELAHKDFVSALPKLDHBCAEGENISVGRQOLVCL 1440
DB 1381 SSGSLRMLNLDPPSOYSDDEWVTSLELAHKDFVSALPKLDHBCAEGENISVGRQOLVCL 1440
QY 1441 ARALLRKTIXLLVDEATAVADLETDDLIOSTIRTOFEDCTVLTAAHLANTIMDTRYIVL 1500
DB 1441 ARALLRKTIXLLVDEATAVADLETDDLIOSTIRTOFEDCTVLTAAHLANTIMDTRYIVL 1500

```

```

QY 1501 DKGEIOEYGA PSDLLQORGLFYSMAXDAGLY 1531
DB 1501 DKGEIOEYGA PSDLLQORGLFYSMAXDAGLY 1531

RESULT 9
AA78873
ID AA78873 standard; protein: 1531 AA.
XX
AC AA78873;
XX
DT 19-MAY-2000 (first entry)
XX
DE Multidrug resistance protein (MRP) natural variant amino acid sequence.
XX
KW Multidrug resistance protein; MRP; human; anthracycline; Vinca alkaloid;
KW epipodophyllotoxin; cancer; leukaemia.
XX
OS Homo sapiens.
XX
EN US6025473-A.
XX
PD 15-FEB-2000.
XX
PE 05-JUN-1995; 95US-00461384.
XX
PR 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Cole SPC, Deeley RG;
XX
DR WPI: 2000-181838/16.
XX
DR N-PSDB; AA290193.
XX
PT Isolated protein conferring multidrug resistance, to at least two drugs
PT selected from anthracyclines, epipodophyllotoxins and Vinca alkaloids, on
PT a drug sensitive mammalian cell.
XX
PS Claim 10; Col 79-88; 78pp; English.
XX
CC This sequence represents a human multidrug resistance protein (MRP)
CC natural variant amino acid sequence. The human MRP confers multidrug
CC resistance, including resistance to at least two drugs selected from
CC anthracyclines, epipodophyllotoxins and Vinca alkaloids, on a drug
CC sensitive mammalian cell, when the protein is expressed in the cell. The
CC multidrug resistance is not substantially reversed by chemosensitizers
CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP
CC protein sequence can be used to generate antibodies against MRP. The MRP
CC protein and nucleotide sequences can be used in compositions which are
CC used to treat patients with tumors displaying multidrug resistance. The
CC compositions and methods of the invention can be used particularly to
CC treat breast cancer, leukaemia, fibrosarcomas, cervical cancer, gliomas,
CC thymomas, neuroblastomas, and lung cancer. Antibodies directed against
CC MRP can be used to inhibit the multidrug resistance of a multidrug
CC resistant cell
XX
SQ Sequence 1531 AA:
SQ
Query Match 80.7%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSPIDMDMNTVNTSNPDKCFQNTYLVVPCGYLMACPFYLYSRH 60
DB 1 MALRGFCSADGSPIDMDMNTVNTSNPDKCFQNTYLVVPCGYLMACPFYLYSRH 60
QY 61 DRGYIOMTPLNKTXTALGLFLMTVCWADLFYSFWEBSRGIFLAPVFLVSPTLLGITTLA 120
DB 61 DRGYIOMTPLNKTXTALGLFLMTVCWADLFYSFWEBSRGIFLAPVFLVSPTLLGITTLA 120

```

QY 121 TPLIOLERRRGVOSSGIMLTFFMLVALVCALAIILRSKIMTALKEDAQVDLFEDITFFYYFS 180
Db 121 TPLIOLERRRGVOSSGIMLTFFMLVALVCALAIILRSKIMTALKEDAQVDLFEDITFFYYFS 180
QY 181 LLLIQLVLSCSDSDSPFSETIHDNCPRESSASFLSITFWMTTGLIVRGYRQPLGSD 240
Db 181 LLLIQLVLSCSDSDSPFSETIHDNCPRESSASFLSITFWMTTGLIVRGYRQPLGSD 240
QY 241 LMSLNKEDTSBOVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPKRESSKVDANEVEAL 300
Db 241 LMSLNKEDTSBOVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPKRESSKVDANEVEAL 300
QY 301 IVKSPKEMNBSLFKVLYKTGPYFLMSFFPKAIHDLMMFSGPOLIKLLIKFVNDTRAPD 360
Db 301 IVKSPKEMNBSLFKVLYKTGPYFLMSFFPKAIHDLMMFSGPOLIKLLIKFVNDTRAPD 360
QY 361 WQGFYTVVLVFTVYCLQTLVHOYFHICFVSGMIRKTAIVGAVYRKALVTNSARKSTV 420
Db 361 WQGFYTVVLVFTVYCLQTLVHOYFHICFVSGMIRKTAIVGAVYRKALVTNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNGPSVLGAVAVMLAMPVN 480
Db 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNGPSVLGAVAVMLAMPVN 480
QY 481 AVMAAMKTQYVAMKSKONRIKLMNEILNGIKYLKLYAMELAFKDYALROBELKYLK 540
Db 481 AVMAAMKTQYVAMKSKONRIKLMNEILNGIKYLKLYAMELAFKDYALROBELKYLK 540
QY 541 KSAIILSVAGFTTWTCTPFLVALCTPAVYVITDENNIIDAOQAPVSLALFNILRPPNLI 600
Db 541 KSAIILSVAGFTTWTCTPFLVALCTPAVYVITDENNIIDAOQAPVSLALFNILRPPNLI 600
QY 601 MVISIYQASVSLKRLRIFLSHEELEPDSIBRRPKYDGGGNSITVRNATFTMARSDPT 660
Db 601 MVISIYQASVSLKRLRIFLSHEELEPDSIBRRPKYDGGGNSITVRNATFTMARSDPT 660
QY 661 LMGITFSIPEGALVAVVQVCGKSSLSALLAEMDKVEGVAIKGSVAVYVPOQAMIOND 720
Db 661 LMGITFSIPEGALVAVVQVCGKSSLSALLAEMDKVEGVAIKGSVAVYVPOQAMIOND 720
QY 721 SLRENILFGCOLBEPRYSVIOACALLPDLILPSGDBTEIGEKVNISSGQOKRVSLAR 780
Db 721 SLRENILFGCOLBEPRYSVIOACALLPDLILPSGDBTEIGEKVNISSGQOKRVSLAR 780
QY 781 AVYNSADIYLPDDPLSAVDAAVGHKIFENVIGPKCMLKNKTRILVTHSMSTLPQVDVITV 840
Db 781 AVYNSADIYLPDDPLSAVDAAVGHKIFENVIGPKCMLKNKTRILVTHSMSTLPQVDVITV 840
QY 841 MSGKISEMGSYOELLARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKOMENG 900
Db 841 MSGKISEMGSYOELLARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKOMENG 900
QY 901 LVYTSAGKOLOROLUSSSSSYSGDISRHNSSTBELOKAEAKKEETKLMLEADYKAOTGYKL 960
Db 901 LVYTSAGKOLOROLUSSSSSYSGDISRHNSSTBELOKAEAKKEETKLMLEADYKAOTGYKL 960
QY 961 SVYVDYMAIGLFTSFLSIFLPMCHVSAIASNYWLSLMTDPIYNGOETHKVPLSYVG 1020
Db 961 SVYVDYMAIGLFTSFLSIFLPMCHVSAIASNYWLSLMTDPIYNGOETHKVPLSYVG 1020
QY 1021 ALGISOGIAVFGYSMAVASIGGILASRCLHVDLHSLILRSPMSFFERTSGNULVNFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVASIGGILASRCLHVDLHSLILRSPMSFFERTSGNULVNFSEKL 1080
QY 1081 DTVDSMIPEVIKMFNGSLFNVIACIVILATPIAIIIPGLIYFFVQRPYVASRQL 1140
Db 1081 DTVDSMIPEVIKMFNGSLFNVIACIVILATPIAIIIPGLIYFFVQRPYVASRQL 1140
QY 1141 KRLSVSVSPYVSHNFTLLGVSVIRAFEEORFTHQDLDKVDENQKAYYSIVANRWLA 1200
Db 1141 KRLSVSVSPYVSHNFTLLGVSVIRAFEEORFTHQDLDKVDENQKAYYSIVANRWLA 1200

QY 1201 VRLCEVNCIVLPAALFAVUSRHSLSAGLVGSYSLSQVTTYIANMLVRMSSEMETNIVA 1260
Db 1201 VRLCEVNCIVLPAALFAVUSRHSLSAGLVGSYSLSQVTTYIANMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSWPQYGRVFRNYCLRYEDDLFPVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPMOIOETAPSSWPQYGRVFRNYCLRYEDDLFPVLRHINVTINGG 1320
QY 1321 EKVGIVRTGAKSSLLTGLFRINESAGEIIIDGINIAKIGHDLRFXITIIIPQDPVL 1380
Db 1321 EKVGIVRTGAKSSLLTGLFRINESAGEIIIDGINIAKIGHDLRFXITIIIPQDPVL 1380
QY 1381 SGSLRKNMLDPPSQYSDPEVWTSLELALXDFVSALPKDLHCEKGGESNLSVGQROLVCL 1440
Db 1381 SGSLRKNMLDPPSQYSDPEVWTSLELALXDFVSALPKDLHCEKGGESNLSVGQROLVCL 1440
QY 1441 ABALIRKTKILVDEAFAAVDLETDLIQSTIRTOFEDCTVLTARHLNTIMDYTRIVL 1500
Db 1441 ABALIRKTKILVDEAFAAVDLETDLIQSTIRTOFEDCTVLTARHLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGAPSDILOQRGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDILOQRGLFYSMAKDAGLV 1531

RESULT 10
ABG61810
ID ABG61810 standard; protein; 1531 AA.
XX
AC ABG61810;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #11.
XX
KM Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN MO200230268-A2.
XX
PD 18-APR-2002.
XX
PE 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733388.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 06-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
DR WP1; 2002-471335/50.
DR N-PSDB; ABK92125.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
PS Claim 27; Page 309; 436pp; English.
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-

CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
SQ Sequence 1531 AA;
Query Match 80.7%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRGFSGADSGDMLDMNVTWNTSNPDTKCFONTYLVWPCFYLMACPFYFLYLSRH 60
Db 1 MALRGFSGADSGDMLDMNVTWNTSNPDTKCFONTYLVWPCFYLMACPFYFLYLSRH 60
Qy 61 DRGYIOMTPINKTKTALGFLMIYCMADLFYSFMEERSGRIPLAVPLVSPFLIGITLLA 120
Db 61 DRGYIOMTPINKTKTALGFLMIYCMADLFYSFMEERSGRIPLAVPLVSPFLIGITLLA 120
Qy 121 TFLIQLERRRGVSSGIMLTFWVALYCALAIIRSKIMTALKEDAOYDFPDITFYVYFS 180
Db 121 TFLIQLERRRGVSSGIMLTFWVALYCALAIIRSKIMTALKEDAOYDFPDITFYVYFS 180
Qy 181 LLLIQLVLSGFSDBSPFSETIHDNPNCPBSSASFLSRITRWITGILVRCYROPLEGS 240
Db 181 LLLIQLVLSGFSDBSPFSETIHDNPNCPBSSASFLSRITRWITGILVRCYROPLEGS 240
Qy 241 LMSLNKEDTSEQVVPVYVWKWKKECAKTRKOPVAVVSSKOPAPCKSSKYDAEEVEAL 300
Db 241 LMSLNKEDTSEQVVPVYVWKWKKECAKTRKOPVAVVSSKOPAPCKSSKYDAEEVEAL 300
Qy 301 IVKSPQKEMNPSELEKVLKTFGPFYFLMSFEKAIHDLMEFSGPOLIKLIFVNDTRAPD 360
Db 301 IVKSPQKEMNPSELEKVLKTFGPFYFLMSFEKAIHDLMEFSGPOLIKLIFVNDTRAPD 360
Qy 361 WQGYFYVLLFVTRACLOTVLHGYFHHICFVSGMRKTAIVIGAVTRKALVITNSARKSTV 420
Db 361 WQGYFYVLLFVTRACLOTVLHGYFHHICFVSGMRKTAIVIGAVTRKALVITNSARKSTV 420
Qy 421 GEIVNLMSVDAPRMDLATYINMWSAPLOYIYALYILMLNIGSVYLAGVAVVWVLMYV 480
Db 421 GEIVNLMSVDAPRMDLATYINMWSAPLOYIYALYILMLNIGSVYLAGVAVVWVLMYV 480
Qy 481 AVAMAKTKTYOVAMKSKDNRIKLMNEILINGIKVLYLAMELAFKDYLAIROBELKVLK 540
Db 481 AVAMAKTKTYOVAMKSKDNRIKLMNEILINGIKVLYLAMELAFKDYLAIROBELKVLK 540
Qy 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVITDENNNIIDAOFAVSLAFNILLRPNLILP 600
Db 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVITDENNNIIDAOFAVSLAFNILLRPNLILP 600
Qy 601 MVISSIVQASVSLKRLIFLSHEELPDSIERRPVKOGGGNSTTVNNAFTFARSPPT 660
Db 601 MVISSIVQASVSLKRLIFLSHEELPDSIERRPVKOGGGNSTTVNNAFTFARSPPT 660
Qy 661 LMGITFESIPGALVAVVGVGCGKSSLSALLAMEMDKVEGHVALKGSVAVVPQAWIIOND 720
Db 661 LMGITFESIPGALVAVVGVGCGKSSLSALLAMEMDKVEGHVALKGSVAVVPQAWIIOND 720
Qy 721 SLAENILFGQLEBPYRSVIQACALLPDLIELPSGRTIIGEGVNLSSGQKORVSLAR 780
Db 721 SLAENILFGQLEBPYRSVIQACALLPDLIELPSGRTIIGEGVNLSSGQKORVSLAR 780
Qy 781 AAVSNADIYFDDPLSAVDAHVGKHIPENYIGPKMLKNKTRILVTHSMYSLPQVDYIIV 840
Db 781 AAVSNADIYFDDPLSAVDAHVGKHIPENYIGPKMLKNKTRILVTHSMYSLPQVDYIIV 840

Db 781 AAVSNADIYFDDPLSAVDAHVGKHIPENYIGPKMLKNKTRILVTHSMYSLPQVDYIIV 840
Qy 841 MSGGKISEMSYQELLARDGAFAEFLRTTYASTBOBAENGVTGSGPGKEAKOMENGM 900
Db 841 MSGGKISEMSYQELLARDGAFAEFLRTTYASTBOBAENGVTGSGPGKEAKOMENGM 900
Qy 901 LMTDSACKOLOROLSSSSSYSGDISRRHNSHTAELOKKAKEETWKLMEDAKOTGVKL 960
Db 901 LMTDSACKOLOROLSSSSSYSGDISRRHNSHTAELOKKAKEETWKLMEDAKOTGVKL 960
Qy 961 SYVWDYKAIQGLFISFLSIFLPMCNHVSALASNYMLMTDDPIVNGTOEHTKRLSVYG 1020
Db 961 SYVWDYKAIQGLFISFLSIFLPMCNHVSALASNYMLMTDDPIVNGTOEHTKRLSVYG 1020
Qy 1021 ALGISQIAVFGYSMAVSGIILASRCLAVDLHSIIRSPMSFEERTPSGNLVNRFSEKL 1080
Db 1021 ALGISQIAVFGYSMAVSGIILASRCLAVDLHSIIRSPMSFEERTPSGNLVNRFSEKL 1080
Qy 1081 DTWDSMIPEVIAKPMFMSLFNVIGACIYILATPIAIIIPPLGLIYFVQRFVASSROL 1140
Db 1081 DTWDSMIPEVIAKPMFMSLFNVIGACIYILATPIAIIIPPLGLIYFVQRFVASSROL 1140
Qy 1141 KRLESYSRSFVYSHFNETLIGSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANWLA 1200
Db 1141 KRLESYSRSFVYSHFNETLIGSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANWLA 1200
Qy 1201 VRLCVCNCIYLPALPAVISRHSLSAGLVGSVSYSLQVTTYLNLVMSSEMETNIVA 1260
Db 1201 VRLCVCNCIYLPALPAVISRHSLSAGLVGSVSYSLQVTTYLNLVMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKAPMPQIOETAPSSWPQYGRVFEFRNYCIRYEDDLFVLRHINVTING 1320
Db 1261 VERLKEYSETEKAPMPQIOETAPSSWPQYGRVFEFRNYCIRYEDDLFVLRHINVTING 1320
Qy 1321 EKYGIYGRTAGKSSLTGLGFRINESAGEEIIIDGINIAKIGLHDLAFKTIIPQDPVLF 1380
Db 1321 EKYGIYGRTAGKSSLTGLGFRINESAGEEIIIDGINIAKIGLHDLAFKTIIPQDPVLF 1380
Qy 1381 SGLSRNMLDPSQYSDDEVWTSLELAHLKDFVSLPDKLDHECAGEGENLSVGRQVLVCL 1440
Db 1381 SGLSRNMLDPSQYSDDEVWTSLELAHLKDFVSLPDKLDHECAGEGENLSVGRQVLVCL 1440
Qy 1441 ARALLRKTILVUDEATAVNDLSTDLIOSTIRTOFEDCTVLITARLNTIMDYTRIVL 1500
Db 1441 ARALLRKTILVUDEATAVNDLSTDLIOSTIRTOFEDCTVLITARLNTIMDYTRIVL 1500
Qy 1501 DKGEIOEYGA PSDLLOORGLFYGMADAGLV 1531
Db 1501 DKGEIOEYGA PSDLLOORGLFYGMADAGLV 1531
RESULT 11
ID ABM35012 standard; protein; 1531 AA.
XX ABM35012;
AC 08-OCT-2003 (first entry)
DT 08-OCT-2003 (first entry)
XX
DE Cancer based on CYP3A5 related protein SEQ ID NO:678.
XX
KM Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;
KM cytochrome p450; subfamily 11A; nifedipine oxidase; polypeptide 5;
XX cytochrome.
XX
OS Unidentified.
XX
PN W02003013534-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002MO-EP008219.
XX

PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kern R;
XX
DR WPI; 2003-268144/26.
XX
XX New use of irinotecan for preparation of compositions for treating cancer
PT in subject having genome with variant allele comprising cytochrome p450,
PT subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.
XX
XX
PS Disclosure; SEQ ID NO 678; 86pp; English.
XX
XX The present invention describes the use of irinotecan (I) or its
CC derivative for the preparation of a pharmaceutical composition for
CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC cancer, or malignant glioma in a subject having a genome with a variant
CC allele, which comprises a cytochrome p450, subfamily IIIA (nifedipine
CC oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have
CC cytostatic activity. The therapeutic applications of (I) is improved,
CC since it is possible to individually treat a subject with an appropriate
CC dosage and/or an appropriate derivative of (I). Therefore, undesirable,
CC harmful or toxic effects are efficiently avoided. Unnecessary and
CC potentially harmful treatment of those subjects who do not respond to the
CC treatment with substances (nonresponders), as well as the development of
CC drug resistances due to suboptimal drug dosing can be avoided. AC62200
CC to AC622751 and ABM34912 to ABM35013 represent sequences used in the
CC exemplification of the present invention
XX
XX Sequence 1531 AA;
SQ
Query Match 80.7%; Score 7860; DB 6; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCGADGSDPLMDMNTWNTSNPDFTKCFONTLVWVPCFYLMACPPFFLYLSRH 60
DB 1 MALRGFCGADGSDPLMDMNTWNTSNPDFTKCFONTLVWVPCFYLMACPPFFLYLSRH 60
QY 61 DRGTYQMPLNKTATGALGFLIMYVCMADLFYSFWRSGITFLAPFLVSPFLTGTTLLA 120
DB 61 DRGTYQMPLNKTATGALGFLIMYVCMADLFYSFWRSGITFLAPFLVSPFLTGTTLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFFMLVALVCALALIRSKIMTALKEDAOVDFRDIFFYVFS 180
DB 121 TFLIOLERRKGVSSGIMLTFFMLVALVCALALIRSKIMTALKEDAOVDFRDIFFYVFS 180
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCESSASFLSRITFMWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCESSASFLSRITFMWITGLIVRGYRQPLEGSD 240
QY 241 LWSLNKRETSBOVAVLVKMKKECAKTRKOPVKKVYSSKDPACKESSKTDANEVVAL 300
DB 241 LWSLNKRETSBOVAVLVKMKKECAKTRKOPVKKVYSSKDPACKESSKTDANEVVAL 300
QY 301 IVKSPKEMNSPLFVLYKTFGPYFLMSPFFKAIHDLMMFSGPOLIKLIRFVNDTKAPD 360
DB 301 IVKSPKEMNSPLFVLYKTFGPYFLMSPFFKAIHDLMMFSGPOLIKLIRFVNDTKAPD 360
QY 361 WQGYFFYVLVFTVLAQTLLVLAHQYFHCIFVSGMRITKAVIGAIVRKALVITNSARKSSTV 420
DB 361 WQGYFFYVLVFTVLAQTLLVLAHQYFHCIFVSGMRITKAVIGAIVRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVILAYLLMLNIGPSVLAVGVAVMLVMEVN 480
DB 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVILAYLLMLNIGPSVLAVGVAVMLVMEVN 480
QY 481 AVMAKTKTYOVAAHMSKDNRIKLMNEITLNGIKVLKLYAMELAFPKDYLAIRROELKYLK 540
DB 481 AVMAKTKTYOVAAHMSKDNRIKLMNEITLNGIKVLKLYAMELAFPKDYLAIRROELKYLK 540

QY 541 KSAVLSAVGTFTTWCTPFLVALCTFPAVYVITDENNIIDAOQAFSLALFNLRPLNLP 600
DB 541 KSAVLSAVGTFTTWCTPFLVALCTFPAVYVITDENNIIDAOQAFSLALFNLRPLNLP 600
QY 601 MVISIVQASVSLKRLIPLISHEELAPDSIERRPVXOGGNSITVNAATTFMARSDEPT 660
DB 601 MVISIVQASVSLKRLIPLISHEELAPDSIERRPVXOGGNSITVNAATTFMARSDEPT 660
QY 661 LMGITFSPREGALVAVVQVCGKSSLLSALLAEMDKVEGHVALKGSVAVVPOQAMTQND 720
DB 661 LMGITFSPREGALVAVVQVCGKSSLLSALLAEMDKVEGHVALKGSVAVVPOQAMTQND 720
QY 721 SLRENILFGCOLLEPPYRSVYQACALLPDLEILPSGRTEIGEGVNLSSGQKORVSLAR 780
DB 721 SLRENILFGCOLLEPPYRSVYQACALLPDLEILPSGRTEIGEGVNLSSGQKORVSLAR 780
QY 781 AVYSNADITYLPDDPLSAVDAAVGHIFENNYGPKMLKNKRIIVTSHMSVLPQVDVYIV 840
DB 781 AVYSNADITYLPDDPLSAVDAAVGHIFENNYGPKMLKNKRIIVTSHMSVLPQVDVYIV 840
QY 841 MSGGKISEMGSYQELRLARDGAFAELRTYASTBOEODAEENGVTGVSQPKAKOMENGM 900
DB 841 MSGGKISEMGSYQELRLARDGAFAELRTYASTBOEODAEENGVTGVSQPKAKOMENGM 900
QY 901 LVYTSAGKOLOROLSSSSSYSGDISRHNSITAELOKAKKEETWKLMEADKAOQGVKL 960
DB 901 LVYTSAGKOLOROLSSSSSYSGDISRHNSITAELOKAKKEETWKLMEADKAOQGVKL 960
QY 961 SVYDYMKAIGLFPSPISIFLFGMCHVSALASNYMWSLMTDDPIVNGTOEHTKRLSYG 1020
DB 961 SVYDYMKAIGLFPSPISIFLFGMCHVSALASNYMWSLMTDDPIVNGTOEHTKRLSYG 1020
QY 1021 ALGISOGIAVGYSAVSIIGIILASRCLHVDLHLSILRSPSPFERTPSGNLVNRFSKEL 1080
DB 1021 ALGISOGIAVGYSAVSIIGIILASRCLHVDLHLSILRSPSPFERTPSGNLVNRFSKEL 1080
QY 1081 DTVDMSLPEVYKMMGSLFNVTGACIYILATPLAIIIPPLGLIYFFVOFYVASSRQL 1140
DB 1081 DTVDMSLPEVYKMMGSLFNVTGACIYILATPLAIIIPPLGLIYFFVOFYVASSRQL 1140
QY 1141 KRLSVSRSPVYSHPNETLGVSVYIRAFEEGERITHOSDLKVDENOKAYYSIVANRMLA 1200
DB 1141 KRLSVSRSPVYSHPNETLGVSVYIRAFEEGERITHOSDLKVDENOKAYYSIVANRMLA 1200
QY 1201 VRLCEVNGCIYLFALFAVISRHSLSAGLVGLSVYSLOVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNGCIYLFALFAVISRHSLSAGLVGLSVYSLOVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSIEKEAPWQIOETAPPSWPQVGRVEFRNYCLRYRDDLDFVLHINVTINGG 1320
DB 1261 VERLKEYSIEKEAPWQIOETAPPSWPQVGRVEFRNYCLRYRDDLDFVLHINVTINGG 1320
QY 1321 EKVGIIVRTGAGKSLTLGLPRINESAGEIIDIIGINAKIGLHDLRKTITIIODPPLF 1380
DB 1321 EKVGIIVRTGAGKSLTLGLPRINESAGEIIDIIGINAKIGLHDLRKTITIIODPPLF 1380
QY 1381 SGLSRMLNDPFSQYSDEEVMSTLELAHLKDFVSAAPKLDHECAEGENLSVGOQOLVCL 1440
DB 1381 SGLSRMLNDPFSQYSDEEVMSTLELAHLKDFVSAAPKLDHECAEGENLSVGOQOLVCL 1440
QY 1441 ARALLRKTIIIVDEATRAVDELFTDDLIQSTIRFOEDCTVLTIAHRLNTIMDYTRIVL 1500
DB 1441 ARALLRKTIIIVDEATRAVDELFTDDLIQSTIRFOEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGASDILLQORGLFYSNAKDAGLV 1531
DB 1501 DKGEIOEYGASDILLQORGLFYSNAKDAGLV 1531
RESULT 12
ADB20865
ID ADB20865 standard; proteoin; 1531 AA.
XX

AC ADB20865;
XX 20-NOV-2003 (first entry)
XX MRP1 based cancer related protein SEQ ID NO:678.
XX
XX irinotecan; colorectal cancer; cervical cancer; gastric cancer;
XX lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
XX variant allele; multidrug resistance protein 1; MRP1; cytosolic.
XX
XX Unidentified.
XX
XX MO2003013533-A2.
XX
XX 20-FEB-2003.
XX
XX 23-JUL-2002; 2002WO-EP008200.
XX
XX 23-JUL-2001; 2001EP-00117608.
XX 24-MAY-2002; 2002EP-00011710.
XX
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
XX Heinrich G, Korb R;
XX WPI; 2003-354397/33.
XX
XX
XX Use of irinotecan or its derivative for preparation of a pharmaceutical
XX composition for treating cancer in a subject having a genome with a
XX variant allele comprising a multidrug resistance protein 1
XX polynucleotide.
XX
XX Disclosure; SEQ ID NO 678; 100pp; English.
XX
XX
XX The present invention describes a method for the use of irinotecan (I) or
XX its derivative for the preparation of a pharmaceutical composition for
XX treating colorectal, cervical, gastric, lung, ovarian or pancreatic
XX cancer, or malignant glioma in a subject having a genome with a variant
XX allele which comprises a multidrug resistance protein 1 (MRP1)
XX polynucleotide (II). (I) has cytostatic activity. (I) or its derivative
XX can be used for the preparation of a pharmaceutical composition for
XX treating colorectal, cervical, gastric, lung, ovarian or pancreatic
XX cancer, or malignant glioma in a subject, where the subject is a human
XX (preferably African or Asian) or a mouse, where the subject represents
XX a sequence which is used in the exemplification of the present invention.
XX
XX Sequence 1531 AA;
XX
XX
XX Query Match 80.7%; Score 7860; DB 6; Length 1531;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 301 IVKSPQKEMWNSLKFVLYKTRGPFYFLMSFFPKAHLDMFSGPOLIKLLIFVNDTAPD 360
QY 361 WOGYFTVLLFTVACLOTVLHOFYFHICFVSGMRKTAIVGAVRKALVTNSARKSTV 420
DB 361 WOGYFTVLLFTVACLOTVLHOFYFHICFVSGMRKTAIVGAVRKALVTNSARKSTV 420
QY 421 GEIYNLMSVPAQRPMDLATYINMTWSAFLQVITLALYLLMLNGSVLAGVAVMLMPVN 480
DB 421 GEIYNLMSVPAQRPMDLATYINMTWSAFLQVITLALYLLMLNGSVLAGVAVMLMPVN 480
QY 481 AVMAKTKTYOVAMHKSCKNRKIKLMEILNGIKVLYAMWELPKDKVLAIRQELKYLK 540
DB 481 AVMAKTKTYOVAMHKSCKNRKIKLMEILNGIKVLYAMWELPKDKVLAIRQELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNLLDAQAFVSLAFNLRPEPLNLP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNLLDAQAFVSLAFNLRPEPLNLP 600
QY 601 MVISIIQASVSLKRLAIFLSHELEPDSIERRPVKDGGGNSITVRNATFTMARSDPT 660
DB 601 MVISIIQASVSLKRLAIFLSHELEPDSIERRPVKDGGGNSITVRNATFTMARSDPT 660
QY 661 LAGITFSPREGALVAVVGVCGCKSSLSLALAEMLDKEGVAKGSVAAYVPOQAMTQND 720
DB 661 LAGITFSPREGALVAVVGVCGCKSSLSLALAEMLDKEGVAKGSVAAYVPOQAMTQND 720
QY 721 SLRENILFGQLEBPYRSYIQAACALLPDEILPSGDRTEIGEKVNLSSGQOKVSLAR 780
DB 721 SLRENILFGQLEBPYRSYIQAACALLPDEILPSGDRTEIGEKVNLSSGQOKVSLAR 780
QY 781 AVYSNADITYFDDPLSAVDHVGKHIFENYIGPKMLKXKTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYSNADITYFDDPLSAVDHVGKHIFENYIGPKMLKXKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGGKISEMSYQBLKRDGAPAEFLRTYASTBEOBDAENGVTGVSFGPGEAKOMENGM 900
DB 841 MSGGKISEMSYQBLKRDGAPAEFLRTYASTBEOBDAENGVTGVSFGPGEAKOMENGM 900
QY 901 LVYDSACKOLQRLSSSSSYSGDISRRHNSTABLOKAEKKEETWKLMEADKQOTGVKL 960
DB 901 LVYDSACKOLQRLSSSSSYSGDISRRHNSTABLOKAEKKEETWKLMEADKQOTGVKL 960
QY 961 SVYWDYKKAIGLFTSFSLTFLMCNHTSALASNYWLSLMTDDPIVNGTOHTKRLSVYG 1020
DB 961 SVYWDYKKAIGLFTSFSLTFLMCNHTSALASNYWLSLMTDDPIVNGTOHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSGIILASRCLHVDLHSLIRSPMSFEERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGYSMAVSGIILASRCLHVDLHSLIRSPMSFEERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMIPEVIKMFMSGLFVNIAGCIVILLATPAAIIPPLGLIYFVORFYAASSROL 1140
DB 1081 DTVDSMIPEVIKMFMSGLFVNIAGCIVILLATPAAIIPPLGLIYFVORFYAASSROL 1140
QY 1141 KRLBSVRSFVYSHFNETLLGVSIVIRAFEBQERFIHOSDKVDENQRAYVPSIVANWELA 1200
DB 1141 KRLBSVRSFVYSHFNETLLGVSIVIRAFEBQERFIHOSDKVDENQRAYVPSIVANWELA 1200
QY 1201 VRLCEVCNCTVPAALFAVISRHSLSAGLVGSVSLQVTTYANMLVRRMSSEKENTIVA 1260
DB 1201 VRLCEVCNCTVPAALFAVISRHSLSAGLVGSVSLQVTTYANMLVRRMSSEKENTIVA 1260
QY 1261 VERLKEVSETEKAPMCIQETAPSSSPQGVREFRNYCLRYBEDLDFVLRHINVTINGG 1320
DB 1261 VERLKEVSETEKAPMCIQETAPSSSPQGVREFRNYCLRYBEDLDFVLRHINVTINGG 1320
QY 1321 EKVGIVERTGAGSSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVGIVERTGAGSSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
QY 1381 SGSLRMLDPPFSQVDEEVWTSLELAHLKQFVSLAPDKLHCEGEGENLSVQOROLVCL 1440
DB 1381 SGSLRMLDPPFSQVDEEVWTSLELAHLKQFVSLAPDKLHCEGEGENLSVQOROLVCL 1440

Db 1381 SGIIRKMDPEPSQYSDIEWTSLLEIAHKDFVSALPDKLDEHCABGGENLSVGQRQLVCL 1440
QY 1441 ARALIRKTKILVLEATPAVDLETDDLIQSTIRTOFEDCVLTIAHRLNTIMDYTRVVL 1500
Db 1441 ARALIRKTKILVLEATPAVDLETDDLIQSTIRTOFEDCVLTIAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIQEYGAPSDLLQORGLFYSSAKDAGLV 1531
Db 1501 DKGEIQEYGAPSDLLQORGLFYSSAKDAGLV 1531
RESULT 13
ADB87954
ID ADB87954 standard; protein, 1531 AA.
XX ADB87954;
AC ADB87954;
XX 04-DEC-2003 (first entry)
DT Human UGT1A1 protein sequence SEQ ID NO:678.
XX
DE Human UGT1A1 protein sequence SEQ ID NO:678.
XX
KM irinotecan; cancer; UGT1A1; cytostatic; topoisomerase I inhibitor;
KM colorectal cancer; cervical cancer; gastric cancer; lung cancer;
KM ovarian cancer; pancreatic cancer; malignant glioma;
KM uridine diphosphate glycosyltransferase1 member A1.
XX
OS Homo sapiens.
XX
PN WO2003013536-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002WO-EP008217.
XX
PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-289896/28.
XX
PT Use of irinotecan to treat cancer patient by determining if patient has
PT variant alleles of UGT1A1 gene, administering increased/decreased amounts
XX of irinotecan based on increased/decreased levels of UGT1A1 gene product.
XX
PS Disclosure; SEQ ID NO 678; 107bp; English.
XX
SS The invention relates to the novel use of irinotecan to treat a patient
CC suffering from cancer. This involves determining if the patient has one
CC or more variant alleles of the UGT1A1 gene, and if the patient has one
CC or more of such variant alleles, irinotecan is administered in an increased
CC or decreased amount in comparison to the amount that is administered
CC without regard to the patient's alleles in the UGT1A1 gene. The invention
CC has cytostatic activity. A composition of the invention acts as a
CC topoisomerase I inhibitor. The method is useful for treating a patient.
CC an animal e.g. mouse or a human, preferably African or Asian, suffering
CC from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
CC pancreatic cancer or malignant glioma. The present sequence is udes in
CC the exemplification of the invention.
XX
SO Sequence 1531 AA;
Query Match 80.7%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCAGSDSDPLMDNMVMTNTSNPDTKCFONTLVVWVCFYLMACFPFYFLYLSRH 60
Db 1 MALRGFCAGSDSDPLMDNMVMTNTSNPDTKCFONTLVVWVCFYLMACFPFYFLYLSRH 60
QY 61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLIGITTTLLA 120

Db 61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLIGITTTLLA 120
QY 121 TFLIQERRKGVSSGIMLTFWLVAALVCAALAIRSKMTALKEBDAQVDLPEDITFYVYFS 180
Db 121 TFLIQERRKGVSSGIMLTFWLVAALVCAALAIRSKMTALKEBDAQVDLPEDITFYVYFS 180
QY 181 LLLIQVLVSCFSDRSPLFSETIHDNDCPESSASFLSRITFWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCFSDRSPLFSETIHDNDCPESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLYKWKKECATRKQPVVYVSSQDPAPKSSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLYKWKKECATRKQPVVYVSSQDPAPKSSSKVDANEVEAL 300
QY 301 IVKSPKEMNPSPLEKVLKYTFGPFLMSFPFKAHDLMMFSGPOLKLLIKFVNDTPAPD 360
Db 301 IVKSPKEMNPSPLEKVLKYTFGPFLMSFPFKAHDLMMFSGPOLKLLIKFVNDTPAPD 360
QY 361 WQGYFTVLLFVTAQTLVLHQYFHI CPGSGMKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTAQTLVLHQYFHI CPGSGMKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMVSVDQRPMDLATYINMTWSAPLQVITALLYLMLNLGSGVLAGAVVMTLMPVN 480
Db 421 GEIVNLMVSVDQRPMDLATYINMTWSAPLQVITALLYLMLNLGSGVLAGAVVMTLMPVN 480
QY 481 AVMAKTKTQOVAMKSKDNRIKLMENELNGIKVLYKLYAMELAFKDKVLAIROEELKYLK 540
Db 481 AVMAKTKTQOVAMKSKDNRIKLMENELNGIKVLYKLYAMELAFKDKVLAIROEELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIIDAQTAFLVALFNILREPLNLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIIDAQTAFLVALFNILREPLNLP 600
QY 601 MVISIYQASVSLKRLKIFLSHELEPDSIERRVRKQGGGNSITVNNATFTARSDDPT 660
Db 601 MVISIYQASVSLKRLKIFLSHELEPDSIERRVRKQGGGNSITVNNATFTARSDDPT 660
QY 661 LMGITFSPREGALVAVVGVQCGKSSLLSALLAMDKVKGVAIKGSAVYVPOQAMIOND 720
Db 661 LMGITFSPREGALVAVVGVQCGKSSLLSALLAMDKVKGVAIKGSAVYVPOQAMIOND 720
QY 721 SLRENIIFGCLEBPYRVSIVQACALLPDLIELPSGRTIEGKGVNLSGQKQVSLAR 780
Db 721 SLRENIIFGCLEBPYRVSIVQACALLPDLIELPSGRTIEGKGVNLSGQKQVSLAR 780
QY 781 AVYSNADITYLFDDPLSAVDHVGKIFBNVIGPKGMLKNTKRIIVTSHMSYLPOVDYIIV 840
Db 781 AVYSNADITYLFDDPLSAVDHVGKIFBNVIGPKGMLKNTKRIIVTSHMSYLPOVDYIIV 840
QY 841 MSGKISMGSYOELARDGAFAEFLRTYASTBOEADAENGVGVSGKGEAKOMENGM 900
Db 841 MSGKISMGSYOELARDGAFAEFLRTYASTBOEADAENGVGVSGKGEAKOMENGM 900
QY 901 LVTSAGKQLOROLSSSSSYSGDISRHHNSTAELOKAEKKEFTWKLMEADKQGVYKL 960
Db 901 LVTSAGKQLOROLSSSSSYSGDISRHHNSTAELOKAEKKEFTWKLMEADKQGVYKL 960
QY 961 SVYWDYMKAGLFLISFLIFLPMCNHVSALASNYWLSMTDDPIVNGTOEHTKVLASYG 1020
Db 961 SVYWDYMKAGLFLISFLIFLPMCNHVSALASNYWLSMTDDPIVNGTOEHTKVLASYG 1020
QY 1021 ALGISQGIARVGYSMANYSIGIILASRCLHVDLHSIIRSPMSFFERTPSGULVNRFSKL 1080
Db 1021 ALGISQGIARVGYSMANYSIGIILASRCLHVDLHSIIRSPMSFFERTPSGULVNRFSKL 1080
QY 1081 DTVDSMIEPVKIMMGSLFNIVIGACIVILLATPIAAIIPPLGILYFFVQGFYVASSQOL 1140
Db 1081 DTVDSMIEPVKIMMGSLFNIVIGACIVILLATPIAAIIPPLGILYFFVQGFYVASSQOL 1140
QY 1141 KRLSVSRSPYVSHFNETLLGVSVIRAEEOERPIHOSDLKVDENOKAYVYSIVANRWLA 1200

QY 961 SVYWDYKAIGLFLISFLIFLMCNHVSALASNYWLSMTDPIYNGTOEHTKYRLSVYG 1020
DB 961 SVYWDYKAIGLFLISFLIFLMCNHVSALASNYWLSMTDPIYNGTOEHTKYRLSVYG 1020
QY 1021 ALGISOGIAVNGYMANYSIGIILASRCLHVDLHLSILRSPMSFPERPTSSGULVNRFSKEL 1080
DB 1021 ALGISOGIAVNGYMAVSIIGIILASRCLHVDLHLSILRSPMSFPERPTSSGULVNRFSKEL 1080
QY 1081 DTYSMTPEVIMKMGSLFNIVIGACIVLLATPIAIIIPPLGLIYFVQRFVYASSRQL 1140
DB 1081 DTYSMTPEVIMKMGSLFNIVIGACIVLLATPIAIIIPPLGLIYFVQRFVYASSRQL 1140
QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEOERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEOERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLBECVNCIYLPALFPAVISRHSISAGLVGLSVSYSLQVTTYLNMVLRMSSEMTNIVA 1260
DB 1201 VRLBECVNCIYLPALFPAVISRHSISAGLVGLSVSYSLQVTTYLNMVLRMSSEMTNIVA 1260
QY 1261 VERLKESETEKEAPMOIQETAPSPSPQVGRVREFRNYCLRYREDLDVLRHINVTINGG 1320
DB 1261 VERLKESETEKEAPMOIQETAPSPSPQVGRVREFRNYCLRYREDLDVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
QY 1381 SSGIRPMNIDPSPQYSDSEEWMTSLBLAHLKDFVSALPDLDHECAGSGENLSVGQRQVYCL 1440
DB 1381 SSGIRPMNIDPSPQYSDSEEWMTSLBLAHLKDFVSALPDLDHECAGSGENLSVGQRQVYCL 1440
QY 1441 ARALLRKTKIIVLBATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRIVYL 1500
DB 1441 ARALLRKTKIIVLBATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRIVYL 1500
QY 1501 DKGEIOEYGAESPDLLOQRGLFYSMAXDAGLV 1531
DB 1501 DKGEIOEYGAESPDLLOQRGLFYSMAXDAGLV 1531

RESULT 15
ADB92128
ID ADB92128 standard; protein; 1531 AA.
XX
AC ADB92128;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human MDRI related protein sequence SEQ ID NO:678.
XX
KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW multidrug resistance 1; MDRI; cytosolic; human; UGT1A1; MRP1; TOP1.
XX
OS Homo sapiens.
XX
PN WO2003013535-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002WO-EP008220.
XX
PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-342400/32.
XX

PT New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.
XX
PS Disclosure; SEQ ID NO 678; 104pp; English.
XX
CC The invention relates to a novel use of irinotecan or its derivative for
CC the preparation of a pharmaceutical composition for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject having a genome with a variant allele which comprises
CC a multidrug resistance 1 (MDR1) polynucleotide. A composition of the
CC invention has cytostatic activity. The present sequence is used in the
CC exemplification of the invention.
XX
SQ Sequence 1531 AA;
Query Match 80.7%; Score 7860; DB 7; Length 1531;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTVMTNSNPDFTKCFQNTVYLVWPCFYLMACPFYFLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTVMTNSNPDFTKCFQNTVYLVWPCFYLMACPFYFLYSRH 60
QY 61 DRGYIOMTPINKTKTALGFLIMIVCADLFTYSFMRSGRFLAVFLVSPITLIGITTLA 120
DB 61 DRGYIOMTPINKTKTALGFLIMIVCADLFTYSFMRSGRFLAVFLVSPITLIGITTLA 120
QY 121 TELIQLERRKGVSSGIMLTFPWLVALVCALALUSKIMTALKEBAQVDLPFDITFYVYS 180
DB 121 TELIQLERRKGVSSGIMLTFPWLVALVCALALUSKIMTALKEBAQVDLPFDITFYVYS 180
QY 181 LLLIQLVLSGCPDSPLFSETIHDNPNCPRESSASFLSRIFFWITGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLSGCPDSPLFSETIHDNPNCPRESSASFLSRIFFWITGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKKWKKCAKTRKOPVRYVYSSQDAPKSSSKDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKKWKKCAKTRKOPVRYVYSSQDAPKSSSKDANEVEAL 300
QY 301 IVKSPQKEMNPSLTKVLKTFGPFLMSFFPKAHDLMFSGPOLKLLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLTKVLKTFGPFLMSFFPKAHDLMFSGPOLKLLIFVNDTKAPD 360
QY 361 WQGYFTVILFVTACLOTVLVHQYFHLCEVSGMRKTAIVGAVVRKALVITNSARKSTV 420
DB 361 WQGYFTVILFVTACLOTVLVHQYFHLCEVSGMRKTAIVGAVVRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRMDLATYINMTWSAPLOYITALYLLMLNLGSPVLGAVAVMTLVMPVN 480
DB 421 GEIVNLSVDAQRMDLATYINMTWSAPLOYITALYLLMLNLGSPVLGAVAVMTLVMPVN 480
QY 481 AVMAKTKTYOVAMKSKDNRIKIMNELINCIKYLKYAMELAFKDKYLA ROEELKYLK 540
DB 481 AVMAKTKTYOVAMKSKDNRIKIMNELINCIKYLKYAMELAFKDKYLA ROEELKYLK 540
QY 541 KSAVLSAVGFTTWCTPFLVALCTFAVYVITDENNIIDAQAFYSIALFNILRPLNLP 600
DB 541 KSAVLSAVGFTTWCTPFLVALCTFAVYVITDENNIIDAQAFYSIALFNILRPLNLP 600
QY 601 MYISSIVQSVSLKRLRIFLSHEELPDSIRRPVKOGGINSITVRNATFTWASDPT 660
DB 601 MYISSIVQSVSLKRLRIFLSHEELPDSIRRPVKOGGINSITVRNATFTWASDPT 660
QY 661 LINGTFSIPBEGALVAVGQVCGKSSLSALLAEMDKVEGVALKGSAVYVPOQAMIOND 720
DB 661 LINGTFSIPBEGALVAVGQVCGKSSLSALLAEMDKVEGVALKGSAVYVPOQAMIOND 720
QY 721 SLRENIIFGCOLBEPYRSVIOACALLPDLIELPSGDRTEIGEGVNLSGGOKRVSILAR 780
DB 721 SLRENIIFGCOLBEPYRSVIOACALLPDLIELPSGDRTEIGEGVNLSGGOKRVSILAR 780
QY 781 AVYSNADITYLPDDPLASVADAHVGRHIFENYVIGPKGMLKNKTRILVTHSMSYLPQVDVITV 840

Db 781 AVYNSADIYLFDDPLSAVDAAHVGHKHFENVIGPKMLGNKXTRILVTHSMSTLPQVDVITIV 840
QY 841 MSGKISMSGSYOELLARDGAFAEFLRTYASTEOQDAENGVTGVSGPGKEAKOMENG 900
Db 841 MSGKISMSGSYOELLARDGAFAEFLRTYASTEOQDAENGVTGVSGPGKEAKOMENG 900
QY 901 LVTSAGKQLOROLSSSSSGDISRHHNSTAELOKAEKKEETWKLMEADKAOTGOVKL 960
Db 901 LVTSAGKQLOROLSSSSSGDISRHHNSTAELOKAEKKEETWKLMEADKAOTGOVKL 960
QY 961 SVYWDYMKALGIFISFLSIFLMCNHVSALASNTWLSMTDPIVNGTOEHTKYRLSVYG 1020
Db 961 SVYWDYMKALGIFISFLSIFLMCNHVSALASNTWLSMTDPIVNGTOEHTKYRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSGIILASRCLHVDLHSLRSFMSFFERTPSGNLVNRFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSGIILASRCLHVDLHSLRSFMSFFERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMTPEVIXMFMSGLFENVIGACIVILLATPIAIIIPPLGLIYFFVQRFVYASSROL 1140
Db 1081 DTVDSMTPEVIXMFMSGLFENVIGACIVILLATPIAIIIPPLGLIYFFVQRFVYASSROL 1140
QY 1141 KRLESVSRSPYSHFNETLGVSVIRAFEOERFIHQSDLVKVENOKAYYPSIVANRWLA 1200
Db 1141 KRLESVSRSPYSHFNETLGVSVIRAFEOERFIHQSDLVKVENOKAYYPSIVANRWLA 1200
QY 1201 VRLECVGNCLVLFALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSEMETNIVA 1260
Db 1201 VRLECVGNCLVLFALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVEFRNYCLAYREDLDPVLPHINVTNGG 1320
Db 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVEFRNYCLAYREDLDPVLPHINVTNGG 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIIPQDPVLF 1380
Db 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIIPQDPVLF 1380
QY 1381 SGIIRKMLDPPSQYSDEEWTSLLELAHKDFVSALPDKLDHECAEGGENLSVGQROLVCL 1440
Db 1381 SGIIRKMLDPPSQYSDEEWTSLLELAHKDFVSALPDKLDHECAEGGENLSVGQROLVCL 1440
QY 1441 ARALLRKTILIVDEATAVULETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKTILIVDEATAVULETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

Search completed: December 15, 2005, 15:17:10
Job time : 180.071 secs

(order) 1000 000 000

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:57:04 ; Search time 37.1218 Seconds
(without alignment)
4901.320 Million cell updates/sec

Title: US-10-665-283-4
Perfect score: 9734
Sequence: 1 MALRGCFSADGSDPLMDMNV.....NTIKVPTPLCTARQLDEDRS 1891

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: PIR.80:*
2: Dlr2:*
3: Dlr3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7849	80.6	1531	1	DVHUR multdrug resistanc
2	4484.5	46.1	1527	2	JE0336 canalicular multir
3	3587	36.9	1545	1	S71841 hyptothetical resistanc
4	3507.5	36.0	1541	1	S71839 canalicular multir
5	3289	33.8	1494	2	E89447 protein P57C12.4 (
6	3225	33.1	1573	2	T21219 hyptothetical prote
7	3195.5	32.8	1502	2	T21216 multdrug resistanc
8	2838	29.2	1515	1	S51863 cadmium resistanc
9	2732	28.1	1398	2	T20434 hyptothetical prote
10	2548	26.2	1478	2	T38712 ABC transporter SP
11	2384.5	24.5	1623	2	T01359 ABC transporter At
12	2356.5	24.2	1622	2	D86428 glutathione-S-conj
13	2326	23.9	1495	2	E86428 probable ABC trans
14	2312	23.8	1144	2	T27408 hyptothetical prote
15	2294.5	23.6	1559	1	S64757 probable membrane
16	2263.5	23.3	1488	2	F86428 probable ABC trans
17	2238.5	23.0	1516	2	F84919 glutathione-conjug
18	2230	22.9	1539	2	T78059 ABC transporter-11
19	2144.5	22.0	1335	2	T00961 hyptothetical prote
20	2144.5	22.0	1514	2	T52080 multi resistanc P
21	2140	22.0	1515	2	T52081 MRP-like ABC trans
22	2131.5	21.9	1490	2	T7840 multi resistanc P
23	2129	21.9	1545	2	T46645 sulfonylurea recep
24	2123	21.8	1153	2	T26883 sulfonylurea recep
25	2109	21.7	1545	2	T42751 sulfonylurea recep
26	2104.5	21.6	1511	2	T42711 sulfonylurea recep
27	2098	21.6	1541	2	T42728 sulfonylurea recep
28	2069.5	21.3	1389	2	T47796 ABC transporter-11
29	2062.5	21.2	1661	2	S64800 probable membrane

30	2037	20.9	1121	2	C87973 protein Y43F8C.12
31	2014.5	20.7	1582	2	A56248 sulfonylurea recep
32	2008	20.6	1592	2	S48933 probable transport
33	1888	19.4	1477	2	S64616 YOR1 protein - yea
34	1884.5	19.4	1037	2	T50518 ABC transporter-11
35	1838	18.9	390	2	S68403 inward rectifier P
36	1827	18.8	390	2	JC4689 inwardly rectifyin
37	1805	18.5	1146	2	F84487 probable ABC trans
38	1781	18.3	390	2	A57616 inward rectifier K
39	1770	18.2	390	2	JC7901 inwardly rectifyin
40	1722	17.7	946	1	JC5667 multdrug resistanc
41	1693	17.4	1548	1	DVLNS multdrug resistanc
42	1687	17.3	1421	2	T34225 hyptothetical prote
43	1606	16.5	1427	2	T20903 hyptothetical prote
44	1575.5	16.2	1427	2	T39219 atp-binding casset
45	1558	16.0	1469	2	T50210 probable ABC trans

ALIGNMENTS

RESULT 1

DVHUR
multidrug resistance protein (cell line H69AR) - human
N:Alternate names: multidrug resistance-associated protein (MRP)
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1993 ; Sequence_revision 05-Dec-1998 #text_change 19-Jan-2001
C:Accession: A44231, A37495
R:Coile, S.P.C.; Bhargava, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.;
Science 258, 1650-1654, 1992
A:Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer
A:Reference number: A44231, MUID:93088080; PMID:1360704
A:Accession: A44231
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'MAPTSGTGMRGIPATPTSPAPTRSSCCGLVFTSGPV', 50-1531 <CO1>
A:Cross-references: UNIPARC:UPI00001746CB; GB:L05628; NID:G1835658
A:Experimental source: small cell lung carcinoma cell line H69AR
A>Note: Sequence extracted from NCBI backbone (NCBI:119851); this sequence has been corrected
R:Coile, S.P.C.; Deeley, R.G.
Science 260, 879, 1993
A:Title: Multidrug resistance-associated protein: sequence correction.
A:Reference number: A37495; MUID:93262415; PMID:8098549
A:Accession: A37495
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-60 <CO2>
A:Cross-references: UNIPARC:UPI00001746CC; GB:L05628; NID:G1835658
A>Note: sequence extracted from NCBI backbone (NCBI:131929)
C:GeneticB:
A:Gene: GDB:MRP
A:Cross-references: GDB:136335; OMIM:158343
A:Map position: 16p13.1-16p13.1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmembrane
F:661-844/Domain: ATP-binding cassette homology <ABC1>
F:678-885/Region: nucleotide-binding motif A (P-loop)
F:1310-1503/Domain: ATP-binding cassette homology <ABC2>
F:1327-1333/Region: nucleotide-binding motif A (P-loop)
F:1450-1454/Region: nucleotide-binding motif B

Query Match 80.6%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGCFSADGSDPLMDMNTANTNSNDFKCFONTLVVPCPYLMACRPFYLYSRH 60
Db 1 MALRGCFSADGSDPLMDMNTANTNSNDFKCFONTLVVPCPYLMACRPFYLYSRH 60
Qy 61 DRGYIOWTPLNKTATAGFLMTVCWADLFYSEFMERSRGIFLAPVFLVSPFTLIGITTLA 120
Db 61 DRGYIOWTPLNKTATAGFLMTVCWADLFYSEFMERSRGIFLAPVFLVSPFTLIGITTLA 120

QY 121 TFLIQLERKRGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLPFDITFYVFS 180
Db 121 TFLIQLERKRGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLPFDITFYVFS 180
QY 181 LLLIQLVLSCSDNSPLFSETIHDNPNCPRESSASFLSITTWITGLIVRGYROPLESD 240
Db 181 LLLIQLVLSCSDNSPLFSETIHDNPNCPRESSASFLSITTWITGLIVRGYROPLESD 240
QY 241 LMSLNKEDTSBOQVAVLVKNNKKECAKTRKOPVKVYVSSKDPAPCKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSBOQVAVLVKNNKKECAKTRKOPVKVYVSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPOKEWNSPLFKVLKYTRGPYFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
Db 301 IVKSPOKEWNSPLFKVLKYTRGPYFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGFYFVLLFVYVYACQLTVLHOYFHIQVSGMRKINAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYFVLLFVYVYACQLTVLHOYFHIQVSGMRKINAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDARFMDLATYINMISAPLOVTLALYLLMLNGPSVLGAVVMYLMVFN 480
Db 421 GEIVNLSVDARFMDLATYINMISAPLOVTLALYLLMLNGPSVLGAVVMYLMVFN 480
QY 481 AVMAKTKTYOVANHKSKDNRIKLNEIILNGIKVLKLYAMBLAFKQVLAIROBELKVLK 540
Db 481 AVMAKTKTYOVANHKSKDNRIKLNEIILNGIKVLKLYAMBLAFKQVLAIROBELKVLK 540
QY 541 KSATYISAVGTFTWCTPPLVALCTPAVYVYTDENNILDAQTRFVSLAFNLIRPPLNLIP 600
Db 541 KSATYISAVGTFTWCTPPLVALCTPAVYVYTDENNILDAQTRFVSLAFNLIRPPLNLIP 600
QY 601 MYISSIVASVSKRLRIFLSHEELEPDSIERRPYKDGGSNTIVRNATFTWASDPT 660
Db 601 MYISSIVASVSKRLRIFLSHEELEPDSIERRPYKDGGSNTIVRNATFTWASDPT 660
QY 661 LINGITFSPREGALVAVGVGCGKSSLSALLAENDKYEGHVALKGSVAYVPOQAMIOND 720
Db 661 LINGITFSPREGALVAVGVGCGKSSLSALLAENDKYEGHVALKGSVAYVPOQAMIOND 720
QY 721 SLRENILFGCCLEBEPYRSVIOACALLPDLLEILPBGDRTEICEKVNLSGGQKQVSLAR 780
Db 721 SLRENILFGCCLEBEPYRSVIOACALLPDLLEILPBGDRTEICEKVNLSGGQKQVSLAR 780
QY 781 AVYSNADLYLPDDPLSAVDANVGHKIFENVIGPKGMKNKRIIVYTHSMSTYLPQDVITV 840
Db 781 AVYSNADLYLPDDPLSAVDANVGHKIFENVIGPKGMKNKRIIVYTHSMSTYLPQDVITV 840
QY 841 MSGKISMGSYOELLARDGAFAEFLRTYASTEOQDAEENGVTGSGGKEAKOMENGM 900
Db 841 MSGKISMGSYOELLARDGAFAEFLRTYASTEOQDAEENGVTGSGGKEAKOMENGM 900
QY 901 LVYDSAGQOLROQLSSSSSYSGDISRHNSHTAELOKAEKKKEETWKLMEADKAOQGYVL 960
Db 901 LVYDSAGQOLROQLSSSSSYSGDISRHNSHTAELOKAEKKKEETWKLMEADKAOQGYVL 960
QY 961 SVYMDYMAKIGLFLSIFLPMCHVSLASNWYLSMTDPIYNGQOEHKTVLSYVG 1020
Db 961 SVYMDYMAKIGLFLSIFLPMCHVSLASNWYLSMTDPIYNGQOEHKTVLSYVG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGLILASRCLVLDLHLSILSPMSGFERTSPSGNLVNFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGLILASRCLVLDLHLSILSPMSGFERTSPSGNLVNFSEKL 1080
QY 1081 DTVDSDIMEVIMKFMGSLFNVIAGCIVILLATPIAIIIPRLGIYFPVQRFYVASSRQL 1140
Db 1081 DTVDSDIMEVIMKFMGSLFNVIAGCIVILLATPIAIIIPRLGIYFPVQRFYVASSRQL 1140
QY 1141 KRLESVSSPYVSSHNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYPSTIVANRWLA 1200
Db 1141 KRLESVSSPYVSSHNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYPSTIVANRWLA 1200

QY 1201 VRLCEVNCIVLPALFAVYSRHSLSAGLVGSYSISLOVTTYINMLVRMSSEMETNIVA 1260
Db 1201 VRLCEVNCIVLPALFAVYSRHSLSAGLVGSYSISLOVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMIOETAPSSWPQVGRVEFRVYCIARYEDLPFVLRHINVTNGG 1320
Db 1261 VERLKEYSETEKAPMIOETAPSSWPQVGRVEFRVYCIARYEDLPFVLRHINVTNGG 1320
QY 1321 EKVGIIVGRTGAGKSLTLGLFRINESAGEIIDDGINIAKIGLHDLRFKTIITIPQDPVL 1380
Db 1321 EKVGIIVGRTGAGKSLTLGLFRINESAGEIIDDGINIAKIGLHDLRFKTIITIPQDPVL 1380
QY 1381 SGLIRNMULDPPSOYSDERWVTSLELAHLKDFVSALPKLHECAEGGENLSVGQROLVCL 1440
Db 1381 SGLIRNMULDPPSOYSDERWVTSLELAHLKDFVSALPKLHECAEGGENLSVGQROLVCL 1440
QY 1441 ARALLRRTKILVDEAPVADLETDLIQSTIRQFEDCTVLTIAHRLNTIMDTRYIVL 1500
Db 1441 ARALLRRTKILVDEAPVADLETDLIQSTIRQFEDCTVLTIAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEYGAAPSDLLQORGLFYSMADAGLV 1531
Db 1501 DKGEIOEYGAAPSDLLQORGLFYSMADAGLV 1531

RESULT 2
JB0336
C:Species: Homo sapiens (man)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JB0336
R:Uchiumi, T.; Hinojosa, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, M.;
Biochem. Biophys. Res. Commun. 252, 103-110, 1998
A:Title: Isolation of a novel human canalicular multispecific organic anion transporter,
t.
A:Reference number: JB0336; MUID:99032812; PMID:9813153
A:Accession: JB0336
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-1527 <UCH>
A:Cross-references: UNIPROT:O15438; UNIPARC:UPI0000169859; GB:AF083552
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: ATP
F:1306/Domain: ATP-binding cassette homology <ABC2>

Query Match 46.1%; Score 4484.5; DB 2; Length 1527;
Best local similarity 56.6%; Pred. No. 5.4e-25;
Matches 871; Conservative 261; Mismatches 312; Indels 35; Gaps 9;

QY 8 SADGSDPLMDNVTWNTSNPDTKCPONTVLVWVPCFYLMACPFYFLYLSRRDRGYIOM 67
Db 7 SGEIGSKRWDSNLSVHNENPDLTPCFQNSILAWPCIIYIMWALPCYLLILRHHCQYIIL 66
QY 68 TPLNKTATGALFLIMYCMADLFYSFWERSRGITLAPVFLVSPULLGITTLATPLIQLE 127
Db 67 SHLSKLMKVGLVLLMCVSMADLFYSFGLVGRAPAVFVTPPLVVGVTMLATLILQYE 126
QY 128 RRVQSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLPFDITFYVFSLLIQV 187
Db 127 RLGQSSGVLITFWLCVCAIVPFRSKILAAEBEISDPFFTYIHFALVLSALI 186
QY 188 LSCSDSPLFSETIHDNPNCPRESSASFLSITFMWITGLIVRGYROPLESDLSLNKE 247
Db 187 LACREKRPFFSAGVNDPNPERSAGLSLFLFWMTKAIYGRHPLREKDLMSLKEB 246
QY 248 DTSEQVAVLVKNNKKECAKTRKOPVKVYVSSKDPAPCKESSKYDANEVEALIVKSPQK 307
Db 247 DRQGVVQQLLEAWRKQEKOTARH-----KASAAPGK---NAGEDVLLGAPRP 294
QY 308 EMNPSLFLVLYKTGPFYFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPDMQGYVYT 367
Db 295 R-KPSFLKALLATGSSPLISACFKLLIIDLISFNPOLSTILIRPISNPMGSPWVGFLVA 353

368 VLFFVTAQLQTLVHGYFHCFSVSGMRITKAVAGAVRKALVITNSARKSTVEIYVLM 427
354 GLMFLCSMMQSLIIQHYHYIFVTVGKERTGIMGVIRKALVITNSVRSATVEIYVLM 413
428 SVDAQRMDLATYTNMISAPLOVILAYILMLNIGPSVLAVGAVMYLWVAVNVMKMT 487
414 SVDAQRMDLATPNTLMSAPLOIILAYIFLMQNLGSPVLAVGAVMYLILPLNGAVAVNM 473
488 KTVAVAHMKSNDRIKXLMNEITLNGIKYKLYAMELAPKDKYLAIRGEBLKYLKSAVLSA 547
474 RAFCVKMKKDSIKMSEILNIGIKYKLYAMSPSPFKQYEGIRGEBLQILRTAAVLT 533
548 VGTFTWVCTPPLVALCTFAVAVTIDENNILDAQAFVSLAFNLIRPPLNLPVSSIV 607
534 TTTFTWMSRPLVLTILMYVYVDPNNVDAKAPVSVSLPNTLRPLNLPOLISLT 593
608 QASVSLKRLRIFLSHELEPDSIRRPVKGGSITVNRATFTMARSDPPTNGITFS 667
594 QASVSLKRIQOFLSQEBELDPQSVKRTISPG--YALTIHSGFTMAQDLPTLHSLDIQ 650
668 IPEGALVAVGQCCGSSLSALLAEMDKVEGHALKGSVAVYPOQAMIONDSRENTL 727
651 VPKGALVAVGVPVCGKSSIVSALLGEMEKLEGVHMKGSVAVYPOQAMIONCTLOENVL 710
728 FGQLEPYPYRSVIOACALPDLILPSGDRTEIGEGKVNLSGGOKORVSLARAVSNAD 787
711 FGKLNKRKYQOTLEACALLADLEMLPEGDOTIEGKGINLSGGOKORVSLARAVISDAD 770
788 IYLFDDPLSAVDAAVHAKHIFENVIGPKGMLKNKTRILVTHSMYLPQVDVLIYVSGKIS 847
771 IFLDDPLSAVDSHVAKHIFDHVIGPBGVLAKGTRVLTGHISLPQDFITVLADQVS 830
848 EMGYSQELIARBDGFAEFLRTYASTBOEQAENGVTVSGPGEAKOMENKML----- 901
831 EMGYPPLLRONSGFANFLCNVAPDEOQGLHEDMTLEGADEKALLIBETLISNHTDLT 890
902 ----VTDSAGKOLROSLSSSSSGSD-----ISRHNSTAE-LQKAEKKETMKLEAD 951
891 DNDVTVTVVQKQFROKQSLALSS-DGEGGKRPVPRHIGPSKVYVTEKKDGA--LQOEB 947
952 KAQGVKLSVYVDMYKALIGLIFLSLIFLPMCHNSALSNVYLSMTDPIYVNGTOEH 1011
948 KKAIGTELVSFMDYKAVAGLCTTLAICLVVQSSAALIGAVMLSAMTMDAMDSQNN 1007
1012 TKRVLSTYAGLIGISQIYAVFGYSMAVSIIGTILASRCHVDLHSLISPSFPERTPSGN 1071
1008 TSIRLGYVAAIGLIQGFVLMMAAMAAAGGIOAARVILHQAIIHNKIRSPSFFPTTSGR 1067
1072 LVNRFSELEDTVDSDMIEVIMKFMGSLFNVIGACIVILATPILAIIPPLGLIYFFVOR 1131
1068 ILNCFSDKIYVDDVILAPVILMLNLSFPAISLTVLVMASPTLTVVILPLAVLYTLVQR 1127
1132 FYVASSRQOLKRLSESVSRSPVSHNETLIGSVVIRAFEOERFIHQSDLKVDENQAKAYP 1191
1128 FYAATSRQOLKRLSESVSRSPVSHNETLIGSVVIRAFEOERFIHQSDLKVDENQAKAYP 1187
1192 STYANRMLVRLCEVNCIVLPAALFAYISHSLSAGLVGLSVSYLQVTTVYMLVMS 1251
1188 YIISNRMLISGVEFVGCNVLFALFAVITGRSSINPLGVLSVSYLQVTFALNMIRMM 1247
1252 SEMETNVAVBRLEKVESETEKEAPWQIOETAPRPSWQVGRVREPRNCLAVRELDVLR 1311
1248 SDESNSVAVBRVEXSKTELEAPVVEGSRPPGMPRGVEPRNRSVRRPDLVLR 1307
1312 HINVTINGEKVIGRTGAGKSSLTGLFRLNESAGEEIIIDGINAKIGLHDLRKIT 1371
1308 DLSLHVHGEKVGIVRTGAGKSSMTCLFRILIEAAGEIRIDGLVADIGLHVRQLT 1367
1372 IIPQDPVLFSGSLAMNIDPQSYDEDEWYSLELAHKDPYSALPDKLHCECAGGNTS 1431
1368 IIPQDPVLFSGSLAMNIDPQSYDEDEWYSLELAHKDPYSALPDKLHCECAGGNTS 1427
1432 VGGQVLCLARALLRKTKILVLEATVAVDLETDLLQSTIRTFQEDCTVLTIAHRLNTI 1491

1428 VGGQVLCLARALLRKTKILVLEATVAVDLETDLLQSTIRTFQEDCTVLTIAHRLNTI 1487
1492 MDYTRVVLVDKGEIQEYGAPSDILQORGLFYMAKADGL 1530
1488 MDYTRVVLVDKGVAVFSDSPANLIARGIYGAARDAGL 1526
RESULT 3
571841
multidrug resistance protein, canalicular - human
C/Species: Homo sapiens (man)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C/Accession: S71841, S71840
R/Koenig, U.; Keppler, D.
Submitted to the EMBL Data Library, August 1996
A/Reference number: S71841
A/Accession: S71841
A/Molecule type: mRNA
A/Residues: 1-1545 <KOE>
A/Cross-references: UNIPARC:UPI00001746CD; EMBL:X96395; NID:g1507819; PIDN:CA65259.1; F
R/Buechler, M.; Koenig, U.; Brom, M.; Kartenbeck, U.; Spring, H.; Horie, T.; Keppler, D
J. Biol. Chem. 271, 15091-15098, 1996
A/Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
A/Reference number: S71839; MUID:96279006; PMID:8662992
A/Accession: S71840
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1415-1429, 'VP', 1432-1455, 'E', 1457-1545 <BUE>
A/Cross-references: UNIPARC:UPI00001746CE; EMBL:X96395
C/Genetics:
A/Gene: GDB:ABCC2; CMOAT; ABC; MRP2; CMRP; DJS
A/Cross-references: GDB:6089489; OMIM:601107
A/Map position: 10q24-10q24
A/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C/Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
F/654-837/Domain: ATP-binding cassette motif A (P-loop)
F/671-678/Region: nucleotide-binding motif A (P-loop)
F/117-1510/Domain: ATP-binding cassette homology <ABCC2>
F/1334-1341/Region: nucleotide-binding motif A (P-loop)

Query Match 36.94; Score 3587; DB 1; Length 1545;
Best Local Similarity 47.54; Pred. No. 3.7e-234;
Matches 746; Conservative 286; Mismatches 461; Indels 78; Gaps 20;
3 LRFGCSADGSDPLMDNVNTWNTS-----NEDFTKCFQNTLVVWPCFYLMACEPFYFL-- 55
2 LEKFCN-----STFNSSFLDSPADLPLCFQGTLVVWIPGLFVLALPWQLAHV 51
56 YLSRHDRGYIQMTPLNKTALGFLMYVQADLPFYSFWRSGIFLAPVFLVSPITLGI 115
52 YKSRITKRSSTTKLYLAK-QVFGVFL-ILAAIILAVLITEDSQAVVPAVRVTPSLY-L 108
116 TLLATPLIOLERRKGVQSSGIMLT-FWLAVCALALRSKXIMTALXEDAQVDLPDIT 174
109 GTWLVLLIOLYSQWCKQKNSWFLSFWIISILCGTFQFOGTLIRTLQGD-NSNLAVSCL 167
175 FYVYFSLILQVLVSCFSDRSPLFSETHDNPDCPSSASFLSRITFWMITGLIVGYRQ 234
168 FPISTYGFQILILFSAFSEN-----NSSNNPSIASFSSITSYWDSDIILNGYR 219
235 PLGSDLSLKNKEDTBQVVPVLYVKKWKECAKTR-----OPVYVYSSKD 281
220 PLTLEIVMEVDEEMKTKTILVSKETEMKRELQKARALGRORKSQOONSQALPGLNKN 279
282 PAQPKSSKYDANEBEYALIVKSPKEMWNS--LFLVLYTFCGPVFLMSPFFAIDHLM 339
280 QSGSQDALV---EDVEKKKKSGTKDVPKSLMLALPFTTFMVLKSLGLVNDIFT 336
340 FSGPQILKILIKFVNDTKAPDWQGYFYVTLFTYACLQTLVHGYFHCFSVSGMRITKAV 399
337 FVSPQILKILISFASDRDTVMIGVLCATILFPAALIQSFCLQCYFOLCFKLGKVRTAI 396

Oy	1400	IGAVYRRALVITNSARKSTGEIYVMSVDAQSPMDLATTIIMTWSAPLOVILATILYM	455
Oy	400	MAZYKALATLSNLRKKEYITVEGETVNLMSVDAQKMDJTMVNMHMLWSSVLOIVISIFFLW	456
Oy	460	LNLPSVLAGAVALMVLWVPVNAVMAAKKRTYQVAHMSKDNRIKLMNIIINGIKYVLKLYA	519
Db	457	RELQPSVLAGVWVWLVPINHLISTKSKTIQVKNMKDKRLKIWEILSIGIKILKYFA	516
Oy	520	WELAFKDKVLAIRQBELKVLKKSAYLSAVGFTWCTPEFVALCTFAVYVYIDENNILDA	579
Db	517	WEPFRPDQVONLRKKEILKNLALFASQLQCVVIFVFLQFTVLVASVTFVSYYVLVDSNNILDA	576
Oy	580	QTAFVSLATFNILRPFPLNLTLMVVISSIVQASVSLKRLAIRLSHETLEPDSIERRPVQDCG	639
Db	577	QKAPTSITLFTNILRPLPSMLPMMLISSMLQASVSTERLEKYLGGDLDLSAIRHSCNFD-	634
Oy	640	GTNSITVNAFTPMARSDPPTINGITFSIPGALVAVVGVGCGKSSILSLALBMDVE	699
Db	635	--KAMQFSEAFSTMEHDEAIVRDVNDIMAGOLYAVIGPVGSKSSILSMLGEMEVH	692
Oy	700	GHVALIKGSVAVVPQQAWITQNDLSRENLTFGCQLEBPYRSYIQACALLPDLEILPSGDRT	759
Db	693	GHITIKGTTAVVPQOSWIQNGTIKONILFTEFNEKRYQOVLACALLPDLEMLPGGDLA	752
Oy	760	EIGEGVNLSSGGQORVSLARAVYSNADIIYLEDPLSAVDHANGHITENYIGRKMLYN	819
Db	753	EIGEGVNLSSGQKRISLARATYQNDIYLLDPLSAVDHANGHIFNKYLGPGLKKG	812
Oy	820	KTRLIVTHSMGYLPQVDVYIWMSGKISBMSGYOELARDCAFAPFKRTY-ASTEOBODA	878
Db	813	KTRLIVTHSMHFLPQVDVETVLQNGTIVKXSYSLMLKKGEPKXNLTEFLRHGPEBEA	872
Oy	879	EENGVTGVSGBKEAKOM-----ENGMLVYDSAGOLQROLSSSSSYSGDISRH	928
Db	873	-----TVHDSGESEADYGLISSVEBIPEDAASITMRENSPRRLTSSSSSNGRHLKSL	927
Oy	929	NSTALEQAEAKKKEET-----WKLMEADKAQOTGYKLSVYMDYMKALIFLSFSLIFEM	983
Db	928	RNSLSTRVNSLKEBDEELVKQDKIKKEPIEFGKVFISLEYEYDAILGFSIFPILIAFV	987
Oy	984	CNHVSALASNWLSLMT-DDPIVNGTO--EHTKRLSLVYGALGJSOGIAVFGSMASVI	1039
Db	988	MNSVAFISNMLMSLWTSDSKIFNSTDYPASQORDMKRGVGTALGLAQGIFPIIAHMSAF	1047
Oy	1040	GGILASRCLHVDLLHSILRSPMSFPERTPSGNLVNRFSKELDYDSMIDEVIKFMFGSLF	1099
Db	1048	GFVHASNILHKQILNNTILRAPRPFDDTPTQGRIVNRFAGDLSYVDLTPQSLRWITCFL	1107
Oy	1100	NVIGACIYIILATPIAIIIPPLGLITYFVQQRFYVASROLKRLSESRSFVYSHNETL	1159
Db	1108	GIISTLWICATYPFTIIVIPLGIITYSVOMFVYSTRSORLRDLSVTRSPISYHSFSETV	1167
Oy	1160	LGVSYTRAFPEOBERTHOSDKVBNOKAYPSYIANSMLVLRCEVCNCLVLPALAEAV	1219
Db	1168	SGLPVTRAFPEHQRFPLKNEVRAIDNQCVRFSWITSNMLAIRLELVNLTVPFSALMMV	1227
Oy	1220	ISRHSLSAGLVGLSVYSLSQVTTYILNMLVRRSSSEMETNIIAVERLKESESFEKESAPMOIQ	1279
Db	1228	IYRDTLSGDTYGFVLSNMLNTIQTILNMLVRRMSELETNIIAVERITETTYKVENAPM-VT	1286
Oy	1280	ETAPSSSPVOYGRVEFRNYCLIRREDDLPVLRHIVNTTNGSKRGVIGVRTAGKSSLTLG	1339
Db	1287	DKRPPDPMPSPGKGIQFNMYQVRRPELDLVLRGITCDIGSMKIGVGRTAGKSSLTNC	1346
Oy	1340	LFRINESAGEIIIDGINIAKIGLHDLEFKITTIPODVLPSGSLRMLDPEFSQSDREV	1399
Db	1347	LFRILEAAGGQIILIDGVDIASIGLHDLBKKTIIPODPLTSGSLRMLDPPNNNSDEEI	1406
Oy	1400	WTSLEIAHAKDFVSAIPDKLDHECAEGENISVGCORVUCLARALARTKYLIVIDEATRA	1459
Db	1407	WKALEIAHAKFSVASLOGLSHIEGRAGCNLSIGRQLLCLGRALLRKSXKILVIDERARA	1466
Oy	1460	VDLETDDLIQSTIRTOFEDCTVLTIABRLNTIMDYTRVIVLDKGEIQEYGAPSDLLQGR	1519

Db 1467 VDLFTDNLQITTIQNEFAHCTVITTAHRLHTTIDSDKMWLDNGKITIECSSPEELQIPG 1526

Qy 1520 LFYSNAKDAGL 1530

Db 1527 PFYFAKEAGI 1537

RESULT 4

871839

canaliculular multidrug resistance protein - rat

C.Species: Rattus norvegicus (Norway rat)

C.Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004

C.Accession: S71839

R.Buchhalter, M.; Koenig, J.; Brown, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, D. J. Biol. Chem. 271, 15091-15098, 1996

A.Title: cDNA cloning of the hepatocyte canaliculular isoform of the multidrug resistance

A.Reference number: S71839; MUID:96279006; PMID:8662992

A.Accession: S71839

A>Status: Preliminary; nucleic acid sequence not shown

A.Molecule type: mRNA

A.Residues: 1-1541 <BUS>

C.Cross-references: UNIPROT:O63120; UNIPARC:UPI000012578; EMBL:X96393; NID:G1292881; PI

C.Superfamily: human multidrug resistance protein CMOAT2; ATP-binding cassette homology

C.Keywords: ATP, glycoprotein, nucleotide binding, P-loop, transmembrane protein

F.100-124/Domain: transmembrane #status predicted <TM01>

F.127-151/Domain: transmembrane #status predicted <TM02>

F.160-187/Domain: transmembrane #status predicted <TM03>

F.305-329/Domain: transmembrane #status predicted <TM04>

F.354-381/Domain: transmembrane #status predicted <TM05>

F.431-451/Domain: transmembrane #status predicted <TM06>

F.456-476/Domain: transmembrane #status predicted <TM07>

F.536-564/Domain: transmembrane #status predicted <TM08>

F.574-602/Domain: transmembrane #status predicted <TM09>

F.650-833/Domain: ATP-binding cassette homology <ABC1>

F.667-674/Region: nucleotide-binding motif A (P-loop)

F.966-994/Domain: transmembrane #status predicted <TM10>

F.1018-1046/Domain: transmembrane #status predicted <TM11>

F.1104-1132/Domain: transmembrane #status predicted <TM12>

F.1203-1228/Domain: transmembrane #status predicted <TM13>

F.1330-1337/Region: nucleotide-binding motif A (P-loop)

F.16,1007,1010,1011/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 36.0%; Score 3507.5; DB 1; Length 1541;

Best Local Similarity 46.2%; Pred. No. 9.1e-229;

Matches 721; Conservative 301; Mismatches 471; Indels 69; Gaps 21;

Qy 6 FCSADGSDPLMKNNTWNTSNPFTCFQNTVLYVWPCFLVLMCFPP--YFLYLSRHDRG 63

Db 4 FCNS----TFWPLSL--ESPEDLPLCFEQTVLWVWIPGLGMLLAPQLQSVSVSRSTRS 58

Qy 64 YIQMTPLNKTALGFLWIVCWADLFYSFWMSRSGIFLAPVPLVSPFLIGITLLATFL 123

Db 59 SI--TKFYIAKQFVVFLILAILAIDLSALTBDTGQATVPVRYTNP--ILYLCTWLVLVA 115

Qy 124 IQLEBRKGVQSSGIMLT--FWLVALVCAALILRSKIMATLEDDAQVD--FRDITFYVYFSL 182

Db 116 VGHSHQMCCKRKNWPLSLTWILSVLGVGFQFQ--LIRALKDKDSNNMAYSLVLFVSYGFQ 174

Qy 183 LIQLVLSGSDRSPLFSEFTIHDPNCPSSASPLSRITFWWITGLIVRGYQPLBGSPLW 242

Db 175 IYLLILITAFSGPS-----DSTQTPSVASPLSITTFWSDRYDTVLKGYGKPLLEBWW 226

Qy 243 SLNK-----EDTSEQVVPVLVKN-----WKSCAKTRKQPVKVVYS--SKDPAQPKSSK 290

Db 227 DIDEGFKTRSVTSKEFAAATKDLQKARQAFORLQKSORPEATLTHGLNKQSSQDVLV 286

Qy 291 VDAEEVEALIVYSPQKWNPNPLFKVLVYKTFGQYFLMSPFKAIDHLMMPGSPQITKLKI 350

Db 287 LBEAKKSKSKTKTDYPKSW---LIRSLKFTFHVILKSLFKLKLHDLVLFLNPQLKLKI 343

Qy 351 KFNVDIKADWQGYFYTVLLFVTACIQTLVLAHQFYFHIQVSGMIRIKTAVIGAVYRKALVI 410

Db 344 GFVKSNSNVMWGYICALLMEFAVTLIQSFCLQSYOHCFTVGLMCRFTTWMSSITKALLTL 403
QY 411 TNSAKSSSTVGGIIVNLMMSVDAQRFMDLATIYIMMSAPLOVILALAYLMLNLNPGSVLAGV 470
Db 404 SMLAKQYTIIGSTVNLMSVDQKMDATATYMWGLVMSVYQITLSIFPLMRBEGPSIILAGV 463
QY 471 AVMYLMVPNVAVMAMKTKTYOVAMHKSCKNRKLMNEIINGIKVILKLYAMELAFDKVLA 530
Db 464 GMMVLLIPNGVLATKIRINIQVONMKNDKRLKIMMELISGKILKYPAMEPSFOEQVG 523
QY 531 IROBELKYLKKSAYLSAVGTFTWCGTPELVALCTFAVYVTLIDENNILLDAQTFVSLALFN 590
Db 524 IRKKELKMLNRGQLOSLITFILOITPILVSVTSSVYVLDVANAVLNAEKAFSTITLFN 583
QY 591 ILRPLNLNLPWISSIVQASVSLKRLRIPLSHEEIEPDSIERPKYDGGGJTSITVRNAT 650
Db 584 ILRPLSLMLPMTSSIILOASVSDRLERTLGDDDLDTAKIRVSNPD---KAVFSEAS 639
QY 651 FTWASDPEPTLNGITFESIPEGALVAVGQVGGKSSLLSALLAENDKYEGHVALKGSVAY 710
Db 640 FTWMDLEATIGDVMILDKPGQLVAVVGTVSGKSSLSVAMLGEMENYGHITIGSTAY 699
QY 711 VPQOMIQNDSIRENILEFGCOLPEEPIYRSVIOACALLPDLITLPSGDRTEIGEKVINSG 770
Db 700 VPQOSWINGITKONILFGESEYNEKKYQOVLKACALLPDLITLPGDMAEIGEEKINISG 759
QY 771 GOKQSVSLARAVSNADITYLPDDPLSAVDAAHKGKIFENVISGPKMMLKMKRITLVTSHMS 830
Db 760 GOKQSVSLARAVYQADITYLPDDPLSAVDAAHKGKIFENKVGPBNGLAKGRTIRIFTHGH 819
QY 831 YLPQYDVITVINSKGKISEMGSYOELLARDBGAFAEFLRTYAS-TEOBODAEENGVTYVSGP 889
Db 820 FLPGYDEIVVLGKGTILEKGSYRDLDDKKGVPARAWKTFMKHSGPGEATYVNN----- 872
QY 890 GKAKQOMENGLVT-----DSAG-----KOLQQLSSSSSYSGDISRHANSTLELO-- 935
Db 873 DSEADDDDDGILPTMBEIPEDBAASLAMBRENSLRTTLRSRSSHRSRSGKSLNKSIXKV 932
QY 936 ---KBAKKEETWKLMEADKAQTGOVKLSVMDYKALGLFSLFSLFLMGNHVSALAS 992
Db 933 NVLKEKEKEVEGQKLIKKEPVBETGKVKESIYKITYLOAGWMSILFILLPYGLANNAPFGS 992
QY 993 NYWLSLWT--DDEIVNGT--OEHTRVRLSVYGALGISOGIAVFGVSMAVSIIGILLASRCL 1048
Db 993 NLMISAMTSDSDNLNGTNNSSSHRDMRIGVFGLAGLQICLLISTMSIYACRNASKAL 1052
QY 1049 HVDLLHSLILRSMSFERTPSGNLYNRPSKELDTYDSMIPYIKMFMGSLFNVIACIYI 1108
Db 1053 HGOLLTNILIRAPMREFDTPGRIYNRPSGDISIYDDLLPQLRSMMWCFPIAGTLVMI 1112
QY 1109 LLATPIAIIIPGLGIYFVQRFVYVASSROKLRELSVRSRVSYPHEMETLLGVSITAF 1168
Db 1113 CMATPEVPAIIIPLSILYISQVFEVYVATSKÖLRKSDSVTKSSITSHSEETVGLTIRAF 1172
QY 1169 EEOERFIHOSDLKVDENOKAYVPSIVANRMLAVRLCEGNCIYIPALFAVISRHSISAG 1228
Db 1173 EHQORFLAMNEQDINDQKVPFSWITSNRMLAIRLELVGNLVVPSGALLVYIKRTILGD 1232
QY 1229 LVGLSVSYSLQVTTYLANMLVMSGSMETNIYAVERLIKSETKEAPQIOETAPSSWP 1288
Db 1233 VVGFLSNALNITQTLNMLVEMTSEAFENIIVAVERISEYINVENEAPM-VTDKPPADMP 1291
QY 1289 QVGRVREFNYCIRAYBEDLDFVLRHNVNTINGEKYGIIGRTGAGSSITLGLFRINESA 1348
Db 1292 RHGEIQFNNOYVRRPELDLVUKGITCNKIKSGEKVGVGRGAGKSSILNCLFRLBSAG 1351
QY 1349 GEIILIDGINAKIGHDLRFRKTIIPQDPVIFSGSLRNMLDPFSQYSDEEYVTSLELAHL 1408
Db 1352 GOIILIDGIDVASIGLHDLRERLTIIPQDPILFSGSLRNMLDPFNKYSDEEYVRLALELAHL 1411
QY 1409 KDFPSALPDKLDHEAEGENUSVQOROLVCLARLLKTKILVLDKETAADVLETDLI 1468

```
Db      1412 KSPFSGDGLGILSEVTBEGDNLSIGQRDLCLCGAVLRKSLIIVLDENETAVAVDETSBLI 1471
QY      1469 QSTIRTOPEDCVTLTIAHRLNTFYRVIVLADKEIQEYGAPSDLLQQRGLEFSYMAKDA 1528
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1472 QTTIRKESQCVTITIAHRLATIMDSDKRIMVLNDGNKIYEYGSPELLSNRGSFYLMAXEA 1531
QY      1529 GL 1530
       |;
Db      1532 GI 1533

RESULT 5
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C|Accession: E89447
R|Anonymous: The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A|Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A|Reference number: A15000; MUID:99069613; PMID:19851916
A|Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_el
A|Note: published extra appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A|Accession: B89447
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-1494 <STO>
C|Cross-references: UNIPROT:Q20943; UNIPARC:UPI000017801B; GB:chr_X; PIDN:AAA83299.1; P
C|Genetics:
A|Gene: F57C12.4
A|Map position: X
C|Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
```

[illegible]

QY 668 IPEGLVAVVGVGVGGKSSLSALLAEMDKVEGVAIVGVVQVQAMQINDSLREML 727
Db 665 VNRGQLVTVVGVGVGKSSMLQGLMGEKSSGISMSGRCLTVVQVPMQNNLTNRQNT 744
QY 728 FGCQLEEPYRSVIOACALLPDLEILPGSDRIEIEKGVNLSGCGQKQVSLARAVYSNAD 787
Db 745 FQKQDEYFYSVLDACALVYRDLQILPLGDNTEIEKGINLSGCGKAKISLARAVYQNH 804
QY 768 IYLPDPDLSAVDAHAKKILFENVIGPKGMLKNKTRILVTHSKSYLPQVDVITVNSGKIS 847
Db 805 IYLLDDPMSAVDAHAGSQVFGSVIPBEGMLNKRKTRILVTNELSFLKSDLLVNNEGKIE 864
QY 848 EMGSYGELLARGAFAEFL-----RTVASTQEQDAENGVTVSGGCKEAKQEN 898
Db 865 YSGKTDLLM-QCGAEFQULLIECEKEPERREKASADEDESEPGIMIGSDSEFYDD 923
QY 899 GMLVYDSAGKQLOROLSSS--SSYSGDISRHNSYTA---ELQKAPAKEETWKLMEA--- 950
Db 924 DVM-----ASPIIDHVLGTSMSYTVSGILNRRISISTHKKQRRLSTTSHHSITSASTQ 979
QY 951 -----DKAQTGVKLSVYMDYMAKIGLFLISFLSIFLPMCNHVSALASNTWLSMTWDP 1003
Db 980 TROLTGTRVEVGRVKNQDTYYKYFGAMGMSIAVLVFLQMTTSTIFSMGRNLMTLQWSDN 1039
QY 1004 IV----NNGQETHTKVRLSVYGALGISQGIIVGVGYMAVSGIGILLASRCLHVDLHSLRS 1059
Db 1040 AARSGSNTGQIARILRVLVYAGLGFSEIILFTIGMLSLLYGVGSVRMLHAFLKRNLEFRV 1099
QY 1060 PMSFEPRTPSGNLVNRFSEKELTDVDSMTPEVKMFMGSLFNVNIGACIYILATPIAII 1119
Db 1100 PMAFYDTTPFGRILNRIGKDIETVDVLLPENVQFPAOCLLOVSTLIIIMSTVPGIVT 1159
QY 1120 PPLGLIYFVQRFYVASSRQLKRLSVSRSPYVSHFNELLGVSVIRAFEEQERFIHQSD 1179
Db 1160 IPLSVYMLVMKRYIATSRQLKRLSITRSPYSHLSESIQGSATIRAYHVLVDRFCRLE 1219
QY 1180 LKVENQKAAVPSVIYANWMLAVRLECGNCTVLPALFAVVISRHSLSGLVGLSVSYLQ 1239
Db 1220 TKVDSHVQRYLYNVANWMLSVRLFEIGNCTVLPALFAALTRITTSYGVGLSVSYALN 1279
QY 1240 VTYVLMVLVWMSSEMETNIVAVERLKEYSETEKEKAPWQIQE-TAPSSWPQVGRFEPNTY 1298
Db 1280 ITTVNFAVRQITKLETNIVSEVERKEYAETETELKEMKSEBQKRPQWPEBGATVWNTY 1339
QY 1299 CLRTRFEDLPVLRHINTVINGEKKYIVGRTGAKSSITLGIFRINESAGEIITDGINI 1358
Db 1340 SARVAPGMLVVKQNLNVEIKPEHEKGIIVRTGAGKSSVTLSLFRITTEAABEQIVVDGINTL 1399
QY 1359 AKIGLHDLRFKITIIIPQPVVFGSGLSRMNVLDPFGQYSPDEEVTSLSLAHLDFPVSALPDK 1418
Db 1400 AEIGLHDLRSMLTIIPQDPVLFSGTLRFNLPDPNNYSQGDITKTEMLANLEFVYAHNEQ 1459
QY 1419 LDHCEAEGEENISVGORQVLCIARALFKTKIILVDEATAVVDVETDQLQSTRTQFED 1478
Db 1460 LNYITTEGSDNISVGORQVLCIARALFKTKIILVDEATAVVDVSTDLALOKTIREEPAN 1519
QY 1479 CTVLTIAHRLNTINDYTRVIVLVDKGEIQEYQAPSDLLQOQGL-FYSMAKQAGLV 1531
Db 1520 ATVTIARLNTINDYRRIIVLNDGKVGEEFSPANLISNRSEFYMAKRAGLI 1573

RESULT 7
T42216
multidrug resistance-associated protein homolog MLP-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T42216
R:Hitoshashi, T.; Suzuki, H.; Ito, K.; Ogawa, K.; Kume, K.; Shimizu, T.; Sugiyama, Y.
A:Title: Hepatic expression of multidrug resistance-associated protein-like proteins main
/:Reference number: Z22081; MUID:98279126; PMID:9614210
/:Accession: T42216

[illegible]

Db 887 -RPPPSDAAPVKS-----TSEAQMEPSLDVEVTGLTAGEDSVQYGRVKATYLSYIRA 940
Qy 970 IGLISFLSIFLFMCNHNVSALASNYWLSIMTDDPIVNGTOEHTKRLSVYALGISQGIA 1029
Db 941 VGTFLCTYTLTLFLCQYVASFCQGWLSIMDDPVDVKQKHSALRGSIIFGLGQLQIG 1000
Qy 1030 VFGYSMAVSIGGIIASRCLHVDLHLSILRSPMSFEPTPSGNLVNRFSEKLDYDSMPE 1089
Db 1001 LFAAMAAVFLGGARASCLFRSLIAMDVARSPGFEPTPVGNILNRFSEKTDIVDVDPD 1060
Qy 1090 VIKRPMGSLFENVIGACIYIILATPIAIIIPPLGIYFVQRFYVASSRQLKRLSEVRS 1149
Db 1061 KMRITLTYAFGLVEGLAVSMATPLAIVAILPLMLYAGFOSLYVATCCQLRLSESAYS 1120
Qy 1150 PVSHPFNFTLLGVSYIRAFEEQERFIHOSDLKVDENQKAYPSIVANFMLVRLCEVNC 1209
Db 1121 SVCSHIAETFGSQSVYRFAQOGPPTAQHDALMDENQGISPPRLVADKMLANLELLGNG 1180
Qy 1210 IVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYINMLVMSSEMETNIVAVERLKEYSE 1269
Db 1181 LVFVAATCAVLTKAHLSAGLAFSVSALQVOTLQMTVRSWTDLENSMVAVERVQDYVH 1240
Qy 1270 TEKAPMOIOETAPSSSPQVGRVFRVYCLRYREDLDFVLRHIVTNGEKVIGVIRT 1329
Db 1241 TPKEAPMPLPSSAQPPLPCCGQIEFRDFGLRHPRLPMVAVGLSKTHAGEKVGIVERT 1300
Qy 1330 GAGKSSLTGLPRINESAGEIIFIDGINIAKIGLHDLFKITIIIPQDPEVLFSGSLRMND 1389
Db 1301 GAGKSSLTGLWGLRLRDEAREGGIWDGVPITDMGLHTLASRTIITIQDPEVLFSGSLRMND 1360
Qy 1390 PFSQYSDERVWTSLELAHKDFVSALPDKLDECAEGEBNISVGRQIVCLARALLRTK 1449
Db 1361 LLOENTWDGIIWALETYVTLKAFVTSIPGLOIYECGCGDDLSVGQKQLICLARALLRTK 1420
Qy 1450 ILIVDEAFAAVALDEITDIQSTIRQFEDCTYTLTAHLNLTIMDTRYIVLIDKEIOYRG 1509
Db 1421 ILIIDEAFAAVALDPGTEIOMQALERWFAQCTVLLIAHRLRSNCARVALVMDGQVASEG 1480
Qy 1510 APSDLQOQGLFYSMAXDAGL 1530
Db 1481 SPAQLIAQKGLFYRLAQSGSL 1501

RESULT 8
S51863
cadmium resistance protein YCF1 - Yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YP9302.11c; protein YDR135C
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S51863; A55352; S50233
R:Oliver, K.; Harris, D.
submitted to the EMBL Data Library, February 1995
A:Reference number: S51863
A:Accession: S51863
A:Molecule type: DNA
A:Residues: 1-1515 <OI1>
A:Cross-references: UNIPROT:P39109; UNIPARC:UPI000013A67B; EMBL:Z48179; NID:g665657; PID:R-Szczypka, M.S.; Wemmie, J.A.; Mowse-Rowley, W.S.; Thiele, D.J.
J:Bio1. Chem. 269, 22853-22857, 1994
A:Title: A yeast metal resistance protein similar to human cystic fibrosis transmembrane
A:Reference number: A55352; MUID:94357936; PMID:7521334
A:Accession: A55352
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-679,'R',681-1515 <SZC>
A:Cross-references: UNIPARC:UPI0000168F06; GB:L35237; NID:g556464; PIDN:AAA50353.1; PID:
C:Genetics:
A:Gene: SGD:YCF1; MIPS:YDR135C
A:Cross-references: SGD:S0002542; MIPS:YDR135C
A:Map position: 4R
C:Function:
A:Description: required for cadmium resistance
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein; yeast vacuole
F:287-308/Domain: transmembrane #status predicted <TM>
F:345-366/Domain: transmembrane #status predicted <TM>
F:421-442/Domain: transmembrane #status predicted <TM>
F:446-467/Domain: transmembrane #status predicted <TM>
F:533-555/Domain: transmembrane #status predicted <TM>
F:558-580/Domain: transmembrane #status predicted <TM>
F:566-829/Domain: transmembrane #status predicted <TM>
F:663-670/Region: nucleotide-binding motif A (P-loop)
F:951-972/Domain: transmembrane #status predicted <TM>
F:995-1016/Domain: transmembrane #status predicted <TM>
F:1068-1088/Domain: transmembrane #status predicted <TM>
F:1092-1113/Domain: transmembrane #status predicted <TM>
F:1179-1200/Domain: transmembrane #status predicted <TM>
F:1208-1229/Domain: transmembrane #status predicted <TM>
F:1289-1483/Domain: ATP-binding cassette homology <ABC>
F:1306-1313/Region: nucleotide-binding motif A (P-loop)

Query Match 29.2%; Score 2838; DB 1; Length 1515;
Best Local Similarity 40.3%; Pred. No. 2,1e-183;
Matches 632; Conservative 295; Mismatches 495; Indels 146; Gaps 27;

Qy 28 DFTKCFQNTVLVWVPCFYLMACPEFYF-----LYLSRD-----RGYIQMTPLNKT 75
Db 27 DFTQCFIDGVTL-----NLSAIFMTTIGIRDVLNLCCKKSGIKYRRNWTIVSMALVTL 81
Qy 76 ALGFLMLIVCWADLFYFWEKSRGIFLAPVFLVPTLIGITTLATLQLERRKGVSS 135
Db 82 EIAF-----VSLASNIKEBAEN---FTIVSQVASTMLSLFVALAHMIEYDR--SVVAN 132
Qy 136 GIMTFMWLVALVCLALIRSKIMTALKEADAVDLFRDITFVYRS-----LLIQVLSC 190
Db 133 TVLLFYWLFETFGNFAKLINILRHYTEG-----IWSQGTGFTLTLFQVYTIC 180
Qy 191 FS-----DSSPL-----FSETHDNPCESSASFLSRITFMWITGLIYRGYROPLEG 238
Db 181 ASILLLEALPKYKPLMHQIHQTLIRKKNPDYDSANFSRTFSMSGMLMTGTEKIVYE 240
Qy 239 SDLSLNKEDTSQVQVPLVYGNMKKECAKTRQKQKVYVSSKDAPOPKSSSKVANEEVE 298
Db 241 ADVLKLPNPFSSSELSQLEKNEMENL-----KQKS----- 271
Qy 299 ALIVKSPQKWNPSLFKVLKYTEGPYPLMSFFPAIHDLMMFSGPOLIKLIKTVND-- 355
Db 272 -----NPLSLMAICRTFGSKMLAFAFKAIDHVALFPTOPOLRIILIKETVDYNS 320
Qy 356 -----TKAPDMQGYFTYVTLFTYACQLTYLHOYFHICVSGMRIKTA 398
Db 321 ERQDHSLSQGFENNHPQKLPYVGRFPLAFPMFLVGFQTSVLHQVFLNVPNTGTYIKSA 380
Qy 399 VIGAVYRKALVITTSARKSTVGEIVNLMVSVDAPRFMDLATYINMINSAPLQVILATYLL 458
Db 381 LTAIIYQKSLVNSBASLSTGDIIVNLMSTDVQKDLTQWMLNLSNGPQIIICLYSL 440
Qy 459 WNLGSPVLAVAVVAVVAVPVNAVAMAKTKTYQVAHMKSKONRIKLANEILNGIKVLY 518
Db 441 KYLGNSMWGVIIILIVIMPLNSFLMRLOKLOKQSKOMKQDERRVRVSEILNINIKSLKY 500
Qy 519 AMELAFKQKULAIR-QEELKVLKKSAYLSANGTWTCTPPLVALCTPAVAVVTIDENNIL 577
Db 501 AMEKRYKRLKEVNNKELKNTLKGQCYMAATVSQFNVPVLPSCCTFAAVP-VTEEDAL 559
Qy 578 DAQAFVSLALENLIRPPLNLTLPWVISIVQASVLSKRLRIFLSHEELPEPSIERRPVKD 637
Db 560 TTDLVFPALTLPNLISFLPLMIIPVANSFIASVISIGLFTFTNBEIQPSVQRLPKVK 619
Qy 638 GGGTNSITV-RNATFTWASD--PPTLNGITFSIPEGALVAVAVGQVGGKSSLSALLAE 694
Db 620 NIGDVAIINIGDAFTFLMRKREYKVALKNINFOAKKGNLTICIVGRVSGKTAALSCMLGD 679
Qy 695 MDKVEGHAIVKGSVAVYVQQAWIQNDSARENIIIRGCQLEERYYSVYIOACALLPDLIELP 754
Db 680 LFRVKGFAIVAGSAVAVVQVPMWNGVTKENILGRHRYDAEFYKTIKACALTTIDLAILM 739

1088 DSIRSPVYAQFGALNGLSITIRAYKAYDRMADNGSRMNNIRFTLVNMGANRLGRL 1147
QY 1204 ECVGNCTIVLPALPAVI-----SRHSLSAGLVGLSVSYLQVTTYLNLVMSSEMETNI 1258
DB 1148 ETLGLGLMILWLASRAVWNGAENQOAFASIMGLLSALNITSLTGLVLAISLAENSL 1207
QY 1259 VAVERLKEYSTEEKAPMOIQETAPPSWPOVGRVEFPNYCLRRREDDLFVLRHINVTIN 1318
DB 1208 NAYERVGNVYIEIPPEAPRVIEENRPPPCWPSSGSIKFEDVILRYAPQLPVLHGVSEFRIH 1267
QY 1319 GGEVGVIVGRTGAKSSITLGLFRINESAGEEIIIDGINIAKIGHDRFKITIIPODPV 1378
DB 1268 PTDVGVIVGRTGAGSSILNLPFRIVEVEKRIILDCCVKGFLMDIRKVLGIIPQSPV 1327
QY 1379 LFGSGLRNNLDPFGQYSDDEEVTSLHLNLDKDFVSLPDKLDHECAEGENLSVGORQV 1438
DB 1328 LFGSTVAFNDDPGEHNADIMESERVHLKDTIRRNLGIDAEVSEGENSPVQORQL 1387
QY 1439 CLARALLRKTYILVDEATAVDETDLLIQSTIRTOFEDCTVLTIAHRLNTMDYTRVI 1498
DB 1388 SLRSALLRSKILVDEATAVDTDALIOKTIREEFKSCYMLIIARHLNIIIDCKIL 1447
QY 1499 VLDKGEIOEYCAPSDLIQORGLFYS-MAKDG 1529
DB 1448 VLDGGRVGEFSSPENLISNEGSSPSKMWQSTG 1479

RESULT 12

866428
glutathione S-conjugate transporting ATPase (AtGRP1) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: D86428
R/Theologian: A.; Becker, J.R.; Palm, C.J.; Federgraj, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasey, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltz, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: D86428
A>Status: Preliminary
A:Molecule type: DNA
A/Residues: 1-1622 <STO>
A/Cross-references: UNIPROT:Q9C8G9; UNIPARC:UPI000009CF47; GB:AE005172; NID:g11055814; F
C/Genetics:
A/Map position: 1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 24.2%; Score 2356.5; DB 2; Length 1622;
Best Local Similarity 35.6%; Pred. No. 1e-150;
Matches 569; Conservative 274; Mismatches 564; Indels 193; Gaps 31;

QY 47 WACPPF-YFLVLSRDRYIOMTPNKKTKLGL-----FLMTVCKADLFYSFW 94
DB 8 WYCEPVPNGVWTKVIVAGVATPCADISFVGLSHLVLLICLRMLWITDHRKVDKFC 67
QY 95 ERSR-GIFLA-----PVFLVSPULLGITTL-----LATFLIQLEKRGQS 134
DB 68 LRSKMFSLFLALLAAYATRAEPLFRU---VMRSVDDLOGAGPPPEAEMLVLE-AFAMG 122
QY 135 SGMLTF-----WLVALCALILRSKIMTALKEDAQVDLPDITFYVFSLL 183
DB 123 SALVTVETKTYIHELRFMYFAVIVAGDMVLNLVLSVK-EYSGSKYLILYSVA 181
QY 184 IOLV-----LSCFSDRSPFSEFIH-----PNCPESASLSTITTW 223
DB 182 VQVAFGLILFYVPNLDPYPGYTPVGTENSEDEYEEELPGGENICPERHANLDSIFPSW 241

QY 224 ITGLIVRGYQPLEGSLMSLNKEDTSEQVVPVLVKNWKECATTRQPVKVVSSSDPA 283
DB 242 LNPMTIGSKRPLTEKXVWMLDPTDKTETLMRSPOKSMDELEK----- 285
QY 284 QPKSSVVDANBEVEALIVSPKEMNPSLKYVYKFGFPFLMSFFPKAIDHLMFSGP 343
DB 286 -PK-----PILRLANSLGGRFWMGKWKIGNDSSQVGP 320
QY 344 QILKLILKFNVDTRAPDMOGYFYVLLFVTACLOTVLHQYFHLICVSGMRKTAVIGAV 403
DB 321 LLNELLKSM-QNEPAMIGIYVLSIFGVVGLVLEAQFQVWMEVGRLSALILAAV 379
QY 404 YRKALVTNSARKSSTGVEIVNMSVDAQRPMDLATYINMWSAPLOYILALYLLMLNG 463
DB 380 FRKSLRLTEBERKKFQTKITNLTMTDAESIQICQSLHTMWSAPFRILVALLYOQLG 439
QY 464 PSVLAGVAVWLVMPVNAWAMKTKYQVAMKSKDRIKMLNGLNGIKYVLKLYAMELA 523
DB 440 VASITGALFLVLMFPICOTVIISKTQKLTKBGLQTDKRIGLMNEVLNAMDVKCYAENS 499
QY 524 FKDKVIALROBELKVLKKSAYLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTA 583
DB 500 FQSKVQIVRODELSEWFRKAQLSAPNMFILNISTIVLVTVVSVGFSLGCD-LTPARAF 557
QY 584 VSLAPNLRPLNPLNPMVVISIVQASVSLKRLIFLSHEE-LEPDSIERPYKDGST 641
DB 558 TSLSLFSLRFPPLNIIITQWNAVANSIRLEEVLSSTERVLLPNP---PLEP-GQ 611
QY 642 NSITVRNATFTW-ARSDPPTLNGITFSIPREGALVAVVGVQVCGSSLSLALAEKD-KVE 699
DB 612 PAISIRNGYFWDKSKARPLISNINDIPGLSLVAVVGSTGEGTSLISALMGLPASP 671
QY 700 GHVAIKSVAVVPOQAMIONDSLRENILFCQLEEPYRVSIVQACALLPDLIELPSGDR 759
DB 672 ATVILRGSVAVPQVSMIFNAITVADNLIIFGAPROEKTERIVDTALQHDLELLPGDLT 731
QY 760 EIGKGVNLGGOKORVSLARAVSNADIVLEDDPLSAVDAVHGKIPENYIGKMKLN 819
DB 732 EIGRGVVISGGQKQVRSMAAVVNSDVCILDDPLSLADHNVQOVFEKCI-KRELQ 789
QY 820 KTRILVHNSKYLPOVDVITVMSGKISSEMSYQELLARDOAFAPLRTASTQEQDAE 879
DB 790 TTRVLVNTQHLFSLQVDKILVHGGVKEBGTVELCHSGPLFRLMENACKVEDY--SE 847
QY 880 ENGVTVGSPGKKAQKQENGLVYDSAGKQLOQLSSSSSYGDISRHNSIAELQX--- 936
DB 848 ENGBAEND-QTSYKVPENG-----NANLQQDGI 875
QY 937 -AEAKKEETKLMADRAQTGOVLYSYWDYMKAI-GLFISFLSIFLMCNHVSALASNY 994
DB 876 ETKSKKEGNSVIVKREBERETGVSMKVLERYQNALGAMVVMMLVICVLTQVFRVSSST 935
QY 995 WLSLMTDD--PIVNGQEHTRVLSYVGLAGISQGIAPFGSMASVIGIILASRCIHL 1052
DB 936 WLSMTQSGTKTGCPLEFN---IYALLSFGGVSVTLINSYWLMSILVAAKMDAM 991
QY 1053 LHSITSPMSFEETPPSGNLVNRFSKELDYDSMIPEYIKFMGSLFENVIGACIVILLAT 1112
DB 992 LGSILIRAMVFFQINPLGRIINRPAKMDGIDRIVAAVPMFQSIQAOLSTVILIGVS 1051
QY 1113 PIAAIIIPPLGLIYFVQRFVVASROLKRLSEVSRSPVSHFNETLLGVSVIRAFEOE 1172
DB 1052 TSLMAINPLLVFYGAVLYQNTSREIKRMDSTRSPVVAQFGBALNGLSIRAYKAYD 1111
QY 1173 RFIHOSDLKVDENQAKYIPSIIVANRWLAVERCEGNCIVLPALPAVL-----SRHSLSA 1227
DB 1112 RMAEINGSRMNNIRFTLVNMAANRWLGIRLEVLGGLVMTLASLAVQNGKANQOAYA 1171
QY 1228 GLVGLSVSYLQVTTYLNLVMSSEMETNIIVAVERLKEYSETEKAPMOIQETAPPSW 1287
DB 1172 STMGLLISYALSTISSLTAVILRLASLANSINSEVRYNTYEISEADVLNNRPPGW 1231
QY 1288 PÖVGVREFRNYCLRYREDDLFVLRHINVTINGEKRVGIVGRTGAKSSLTGLFRINESA 1347

Db 1232 PSSSISKFEDVLLRYRPELIPVLHGVSPFLISPMKVGIVGRTGAGKSLNALPRIVELE 1291
 Qy 1348 EGEIIDIINIAKIGLHDLRKITIIIPDPLFGSGLRMLDPPSQSDEEVTWLSLAAH 1407
 Db 1292 KGRLLIDECIDIGRGLMDLRVLTGIIIPQAPVLFSGTVAFNIDPPSEHNDALWESLEBAH 1351
 Qy 1408 LKDFVSLALPDLDHECAEGENLSVGOQLVCLARALLRKTIIVLDEATAVAVLETPDL 1467
 Db 1352 LKOTIRRNPLGLDDEVTEAGENPSVGOQLSLARALLRKTIIVLDEATAVAVRIDVL 1411
 Qy 1468 IQSTIRFOEDCTVLTIAHRLNTIMDYTRVIVLDKGEIQEYGAESDILQGRGFYSMAKD 1527
 Db 1412 IQKTIREFKSCMTLIIAHRLNTIIDCKVLVLDGKGVQEPSSPENILSNESSEFSK--- 1468
 Qy 1528 AGLVGGGGGGLSKRGKIIEEYVLTRLAEDPAEPRYRTRE 1567
 Db 1469 -----MVOSTGTANAEYLSRISITLENK-----RTRE 1493
 RESULT 13
 E86428
 Probable ABC transporter (imported) - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C/Accession: E86428
 R/Theologis, A.; Ecker, J.R.; Palm, C.U.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Ansen, N.F.; Hughes, B.; Hultar, L.
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maltz, R.; Merzajali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: E86428
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1495 <STO>
 A/Cross-references: UNIPROT:Q9C8H0; UNIPARC:UPI00000482FF; GB:AE005172; NID:g11055818; F
 C/Genetics:
 A/Map position: 1
 C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
 Query Match 23.9%; Score 2326; DB 2; Length 1495;
 Best Local Similarity 34.3%; Pred. No. 1e-148;
 Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;
 Qy 2 ALRGFCG--ADGSPPLMDNVTNNTSNDFPKCQNTVLVWVP-----CFY-LMACFP 51
 Db 5 ALMWYCPRVADGF-----WEKAVDGAFCAGAYPCADISLMLVSHFVLLGLCFYRIWIF- 58
 Qy 52 FYFLYSRHRDGYQMTPLNKTALGLFLMIVGMAD-----LFSFWERSRRIFLAPV 105
 Db 59 -----HNTAQIYIVARKYYNCVLG-LACYCVPVRLVLYVGLISLFPMDDEFTDPPFP 110
 Qy 106 FLVSPPLLIGITTLTFLIQLERRKGVQSSGIMLTFLMLVCLALIRSKIMTALKEDA 165
 Db 111 EVASLWAEAFAMFEMVLVIGLETQYKKEPRWYRPFVLYVLVADAVLLDLVLPKXSN 170
 Qy 166 QVLDLFRDIT--FYVYSLLIQLVLSGFSRSPLESETTHDP-----NCP 209
 Db 171 RTALYLISSRCQALFGIILLIYIPRL--DPYGYHI VNNPEPLDNEVYDALRGHEICP 228
 Qy 210 ESSAFSPISRTFMWITGLIVGYRQPLSGSDMLSKNEPSEOVVPLVKNWKKCAKTR 269
 Db 229 ERHAISSRIYFGITPLMQLGIRKPKITEKDVQDLKMDQETETLIKRPQRCTWESRR-- 286
 Qy 270 KQPVKVVYSSKDPQAPKSSKVDANEEVALIVSPQKWNPSLFLKVLKTFGFPYPLMSF 329
 Db 287 -----PK-----PMLRLALNNSISLGRFPLAG 307

Qy 330 FFRAIHLMMSPGQILKLIKFNVDTKAPDMQGYFTVLLFYTACLOTVLHOFHICF 389
 Db 308 IFKIGNDLSQFVGVVILSHILRSKQEGD-PANWGYVYAFIIFPVGTIGLCEAOYFQNW 366
 Qy 390 VSGRIKTAIVAGYRAKALVITNSARKSVTGEIVLMSVDAORFMDLATVYNNMWSAPL 449
 Db 367 RVGRILASTLVAALPHKSLTLTHEARKNFPASGKVTNITTDANLQOISQHLGWSAPF 426
 Qy 450 QVIALYVLMLNIGPSVLGAVAVMTLVPNVAVMAKTKYQVAHMSKNRIKLAMEIL 509
 Db 427 RIIVSMILLQOGLVSLFQSLILFLIPIQTLIISTRKILTEGLOMTKRVGITEIIL 486
 Qy 510 NGIKVLYLAMELAKOKVLAIQEELVKLKSAYLSAVGFTWVCPTPLVACTPAFVY 569
 Db 487 SMDTVACVAMERSFESRIGIRNEBELSMFRKQLSAFNSFILNIPVAVTVVSFGV 546
 Qy 570 TIDENNLDAQTAQVSLALFNILRFPNLIIPWYISSIVQASVSLKRL-RIFLSHELEPD 628
 Db 547 ILGCD-LTPARAFTSLFAVLRFLPMLPNLSQVYVNAVSLQRIEELLSEBRLAQ 604
 Qy 629 SIERRPVKDGQGTNSITVRNATTW-ARSDPTLNGITFSIPGALVAVVQVGGKSSL 687
 Db 605 NPLQF-----GTPAISIKKGYFSWDSKTKPTLSDINLEIPVGTVAIVGTEGKSTL 659
 Qy 688 LSLALLMDKVE-GHVAIKSVAVVPOQAMIQNDLSRENILFGQLEPRTYRTOCAL 746
 Db 660 ISMLGLSLHAETTSVIRKSVAVVPVSVIMFATVARENILFSGDFSEBRYRAIDTAL 719
 Qy 747 LPDLIELPSDRTIEGEGVNLGGOKORVSLARAVVSNADIVLPDLPAVDAAHVKHI 806
 Db 720 QHDDLPLPGDULREIGRGVNISSGQKQKSMARAVVSNSDVYIFDDPLSALDAHVAQV 779
 Qy 807 FENVIGKGMUKKTRILVTHSMSTYLPQVDVIVMSSGKISSEMSYQELLARDAFAFL 866
 Db 780 FDSGM-KDRLRGTRVLVTNQHLPLMDKILIVBSGMKEGTFELSKSGLFLFKM 837
 Qy 867 RTYASTQEODAEENGVTGVSFGKEAKQENMLVYDSAGKOLQROLSSSSYSGDISR 926
 Db 838 ENAGKDATQEVNTN-----DENIKGLPTVTVDS-----ERNLGSYTK----- 876
 Qy 927 HHNSTAELQAEAKKEETWLTMEADKAQOTGVLSYWDYWKAI-GLFISFLIFLPMCN 985
 Db 877 -----QGRKRS-VLIKQERETIGISMVLMRKXBAVGLMVMILLACYLAT 924
 Qy 986 HVSALASNYVLSMTDDPIVNGQEHK-----VRLSVGALGISQGIAGVGSMAVSIG 1040
 Db 925 EYLRVSSSTWLSIWTD-----QSTSKVNSPGFYIVVALLGQGVAVFTMSFWLITS 977
 Qy 1041 GILASRCGLAVDLHSILRSMSFPERTSGNLVNRSKELDTYDSMIPVYIKPMGSLFN 1100
 Db 978 SLHAARLHDAMLSILRAFMLEFHNPTGRVNRSKDGDIDRVANIMNNFMQLO 1037
 Qy 1101 VIGACTVILLATPILAIITPILGLIVFVGRFVAVASROLKLESVRSRVSYPHNETL 1160
 Db 1038 LUSTPALIGVSTISLMAIMPELILFLYAAVLYQSTREKRRLDSVTRSTIYQFGBALN 1097
 Qy 1161 GSVVIRAFEBQERFIHOSDKVDENOKAYYPSIVANRMLAVRLCEVCNCTVFAALFAYI 1220
 Db 1098 GLSSIRAYKAYDMAKINGSMONNIRFLTANNSNRMLIRIETLIGGVIMWLTATFAVL 1157
 Qy 1221 SRHSLG-----AGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVAVERLKEYSETEKAP 1275
 Db 1158 QNGNTNQAGFASFTGMLASTYLNISLGSVROARSRAINSINSVERGVNYIDLSEAT 1217
 Qy 1276 MOIQETAPSPSPQVGVGFERNYCLARYEDLDVFLRHINVTINGKEVGVVIGRTGAGKS 1335
 Db 1218 DILENNRPVCGWSPGSGSIXKEDVHLKRPCLPVLGLITRFPSEKVGAVGRTGAGKS 1277
 Qy 1336 LVLGLFRINESAGEIIIDGINIAKIGLHDLRKITIIIPDPLFGSGLRMLDPPSQS 1395
 Db 1278 MLNALPRIVEVEKRIIMDDQVAKFGLTDRVRLSIIPOSPLFSGTVAFNIDPSEHN 1337

```

QY 1396 DEEWTSLELAHLKDPVSLPDKLDHBCAEGENLSVGQRLVCLARALLKTKIYLVDL 1455
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1338 DAGIMELHRAHAKDIVSRNPFGLDAVECEGENFSVGQRLSLARALLRSKILVLDE 1397
QY 1456 ATAAVDELTDLLIOSTIRTPQEDCTVLTIAARLNTINDYTRVYVLDDGEIOEGAPSDLL 1515
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1398 ATASVDRTDLSLQRTIRFEKSCMTLVIAHRLNTIIDCDKILVLSGQVLEYSPOELL 1457
QY 1516 -OORGLFYMAKDAG 1529
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1458 SRDTSAFRRVHSTG 1472

RESULT 14
272408
hypothetical protein Y75B8A.26 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T27408
R:Barlow, K.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z20361
A:Accession: T27408
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1144 <MIL>
A:Cross-references: UNIPROT:Q9XW61, UNIPARC:UPI000007AA2A; EMBL:AL033514; NID:e1343251;
A:Experimental source: clone Y75B8A
C:GeneticS:
A:Gene: CESP:Y75B8A.26
A:Initons: 63/3; 120/3; 222/3; 271/3; 372/3; 451/1; 636/1; 687/3; 778/3; 953/2; 1042/2;

Query Match 23.8%; Score 2112; DB 2; Length 1144;
Best Local Similarity 42.5%; Pred. No. 5.9e-148;
Matches 491; Conservative 228; Mismatches 381; Indels 54; Gaps 16;

QY 396 KTAIVAGVYRKALVITNSAKSSVTGEIVNLMSVDAQRFMDLATYINNMASAPLOVILAL 455
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 QTVLSNAILHKILNLSPEARSNRTAGEILNHAVDIEIIVHSVPYLQNMWGVPPQVTLAM 68
QY 456 YILMLNIGPSVLAVGVAVVWLVNVPVNAVMAAKTKYQVAHMSKDNRIKLAMNLIINGIYVL 515
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 TMLAITTIGWAMAMAGVCMILFIPNLCTSRPKLSQOKOMKIKERTLSNEMNGIKYV 128
QY 516 KLVAMELAFKDKVLAIRQELKLVKSAVLSAVGFTWCTPFLAALCTPAVYV--TIDE 573
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 KLVAMESFEQINRLRAKVEYKMLRNVCILSRIVDVANASPFVLAIGSTFCYVLMSPDE 188
QY 574 NNILDAQTAFLVSLAFNLRPEPLNLPVVISIYQASVSLKRLRIFLSHEBELPDSIERR 633
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 NG-LTPSVAFVALTIFNQLRQPMRMVANLINTLVQARVSNKRLRQFLNDEME---RKT 243
QY 634 PVKGGGNGNSTITVRNATFTW--ARSDPPLNGITTSIPFGALVAVVGVQVCGCKSSILSL 692
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 244 EVALG---NAIVFKNASLMMKGPONPVPVKDLSAIIRKGOIATVSGVGKSSLSLAVL 300
QY 693 AEMDKGVHVAIKGSVAVPOAMQNDLSRENILFGQLEBPYRSVIOACALLPDEI 752
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 DEWVLLDRKRVGVGSIAVVPQHSWTFNKTIENILFGNELSNYFDQVVGSCQLTDRH 360
QY 753 LPSGDRTEIGKGVNLSSGQKQKRVSLARAVSNADIVFPDPLSAVDAVGHKIFENYIG 812
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 FQGGENTMVGENGITLSSGQKARISLARAVQDKOIVYLDPLSAVDHVRALFDKYG 420
QY 813 PKGMLNKTRLIVTHSNLYLPQVDVITYMSGGKISEMSYQELLARDAFAEFLTYAST 872
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 PDGLRSTKRLVTHNLTQYTKYVDITVYIEBQIYQHSFEDIAVVDGPFGR-LMSSECN 479
QY 873 EQGDADENGVTVSGPGEKAKOMENGMVLVDSAGKQIQRDLSSSSSSGDISRHNSTA 932
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 480 SDEVDADAEASSEASVTVPVPLVENG---DNGAIEKSSQIDRKNSHFSEKSR----- 529
QY 933 ELQRAEAKKEETWKLMEADKAQGTQVKLSVYWDYKAIGLFISFLISFLMCHVSALAS 992

```

```

Db 530 ---KSEKPKQKVEKNV--NVOLGRVKKSVYOLYIKTKMGIFNSAPLIFLAHTVIMR 584
QY 993 NYWLSLWTD-----PIVNGTQ-----HTKVLSSYVALGISQGIAPVFGSM 1035
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 585 SLWLSDMNSNEMAIKATLSSVDYLNSTSSVDGVSVETRLIYAGG--GLMLLLAL 641
QY 1036 A---VSIQILASRCLHVDLLHSLRSPMSFFERTPSGNLVNRSKEIDTVDSMIPEYIK 1092
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 642 AFTVLTIGSLRASYGHSPLIHALLVAPISFPDTPPTGRIINRSRLDVIDK-LQNNIR 700
QY 1093 MEMGSLFNVIAGACVILATPIAAIITPEGLYFFQRTVVASRQLKRLSESRSBPY 1152
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 701 MCTQTLNACMIIVLISISTPIPLVCAAPLILYFVMIYIYIPSRQLKRLESANRSPIL 760
QY 1153 SHNETTLGVSIVRAPEQERFIHQSLDKVDENQKAVPSIVARWMLAVRCEYGNICVL 1212
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 761 STAESIHGASSTIAFPKTEKRTTTLASTNDKFAQCHRLSHMSKMLATRIELIGNICVL 820
QY 1213 FALFAVISR--HSLASGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVAVERLKEYSET 1270
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 821 FASLSATLSKRYFGLTFGMAGLSVYALTTTEVLANICVRSVEIESNIVSVERVNEYOXL 880
QY 1271 EKEAPMOIQETAP--PSSWPOGVREFFRNKYLCTRREDLDFYLRIHNTVINGEKYIGRT 1329
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 881 EPEAPMRIRKSLSENEEKMPYVGKIELDGFMRKRYKNLPVLKNDLKEGGERIGVIGRT 940
QY 1330 GAGSSSLTGLFRINESAEGEIIIDGINIKTIGLHDLRFKTTIIPDPVLPSSGLRMLND 1389
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 941 GSGKSSLTMLALYKRIEBSGTYIKIDVBIDTIGHQLRSKLIIIPQEPVSVSGTLRRFLD 1000
QY 1390 PFSQYSDIEWTSLAHLKDPVSLPDKLDHBCAEGENLSVGQRLVCLARALLKRTK 1449
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1001 PFNYSDQDQIMNGEILQIQKQFQEDDKTIDRYIAEGKKNMSVGERQLLCLCRALLGAR 1060
QY 1450 ILVIDEATPAVDLETDLIOSTIRTPQEDCTVLTIAARLNTINDYTRVYVLDDKEIQEYG 1509
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1061 IVILDEATVADVTVTDGIVQRAIRQHPPOSTTISIAHRLDIYDSDRIVVLDAQVAFED 1120
QY 1510 APSDLLQORGLFYS 1523
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1121 TFSNLLNPDLSLYS 1134

RESULT 15
564757
probable membrane protein YIL015w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein A251; Hypothetical protein L1313
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S64757; S64763; FS0041; S69391; S70560
R:Miooga, T.; Zimmermann, F.K.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64743
A:Accession: S64757
A:Molecule type: DNA
A:Residues: 1-1559 <MIO>
A:Cross-references: UNIPROT:P14772; UNIPARC:UPI0000126A9C; EMBL:Z73120; NID:g1360184; PIR:
A:Experimental source: ECRAN S288C
R:Goffeau, A.; Purnelle, B.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64761
A:Accession: S64763
A:Molecule type: DNA
A:Residues: 1-1559 <GOF>
A:Cross-references: UNIPARC:UPI0000126A9C; EMBL:Z73120; NID:g1360184; PIR:CAA97460.1; PIR:
A:Experimental source: strain S288C
R:Boy-Marcotte, E.; Damak, F.; Camonig, J.; Garreau, H.; Jacquet, M.
Gene 77, 21-30, 1989
A:Title: The C-terminal part of a gene partially homologous to CDC25 gene suppresses the
A:Reference number: FS0041; MUID:89306677; PMID:2545538
A:Accession: FS0041
A:Molecule type: DNA

```

A:Residues: 1-255 <BOY>
A:Cross-references: UNIPARC:UPI00001746CF
A:Note: the authors translated the codon CAG for residue 248 as His
R:Purnelle, B.; Goffeau, A.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of 32 kb on the left arm of yeast chromosome XII reveals 14
mily and a new ABC transporter homologous to the human multidrug resistance protein.
A:Reference number: S69380
A:Accession: S69391
A:Molecule type: DNA
A:Residues: 1-1559 <EUR>
A:Cross-references: UNIPARC:UPI0000126A9C; EMBL:X97560; NID:G1297003; PIDN:CAA6162.1; F
R:Mooga, T.; Zimmermann, F.K.
Yeast 12, 693-708, 1996
A:Title: Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on a 43.7 kb
e conductance regulator protein CTR.
A:Reference number: S70557; MUID:96405918; PMID:8810043
A:Accession: S70560
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1559 <MIM>
A:Cross-references: UNIPARC:UPI0000126A9C; EMBL:X91488; NID:G1495203; PIDN:CAA62776.1; F
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995
C:Genetics:
A:Gene: SGD:BPT1; MIPS:YIL015W
A:Cross-references: SGD:S0003938
A:Map position: 12L
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F:28-44/Domain: transmembrane #status predicted <TM1>
F:143-159/Domain: transmembrane #status predicted <TM2>
F:178-194/Domain: transmembrane #status predicted <TM3>
F:334-350/Domain: transmembrane #status predicted <TM4>
F:421-437/Domain: transmembrane #status predicted <TM5>
F:526-542/Domain: transmembrane #status predicted <TM6>
F:550-566/Domain: transmembrane #status predicted <TM7>
F:654-679/Domain: ATP-binding cassette homology <ABC1>
F:672-679/Region: nucleotide-binding motif A (P-loop)
F:974-990/Domain: transmembrane #status predicted <TM8>
F:1017-1033/Domain: transmembrane #status predicted <TM9>
F:1099-1115/Domain: transmembrane #status predicted <TM10>
F:1118-1134/Domain: transmembrane #status predicted <TM11>
F:1212-1258/Domain: transmembrane #status predicted <TM12>
F:1319-1559/Domain: ATP-binding cassette homology <ABC2>
F:1336-1343/Region: nucleotide-binding motif A (P-loop)

Query Match 23.6%; Score 2294.5; DB 1; Length 1559;
Best Local Similarity 35.1%; Pred. No. 1.5e-146;
Matches 574; Conservative 292; Mismatches 541; Indels 227; Gaps 39;

QY 23 NTSNPDPTKCFQNTVVLVWVCFYIM-ACFPFYLYLSRHDRGYIQMTPLNKTATAGFL-80
DB 23 NALNP-----CFISVISAQAVFLLIGSYQYLMKLYKNNKVPFRKKNPPLTPSKINSRHLT 78
QY 81 -LWIVCW-----ADLFYSPWBSRGIFLAPVLPVSPILLGI-----TITLTA 121
DB 79 HLTNVCQSTLIICELALVSSQSDRYVPLTK-NALYINLIF-----NLGISLPTQYLA 132
QY 122 FLIOLERRKVOSSGIMLTFWLVALYCALAIIIRSKI-----MTALKEDAQV 167
DB 133 F-----KSTFPMGNQLFYVMFOILLQFLIIOORYHSSNRLTYISQGTMIILEVL 185
QY 168 DLFRDITFYVYSLLLIQLVLSQSPDRPLPSETIHDNDPCPSASFLSRTFFWITGL 227
DB 186 -LNSVAIFLY-DLCIFEPT-----NELSEYKKNQWFP-----PVHVSYITFIWMNKL 233
QY 228 IYAGYROPLESGDLMSLTK--EDTSEQVAVLVANMKKECAKTRQKVAVVYSSKQAP 285
DB 234 IYETVR-----NKIKDPNQLPLPVDLNLK----- 259
QY 286 KSSSKVDANEVALLVYKSPQKEMNPSLFLVLYKTFGPYLMGFFFAHIDLMMFGSPQI 345
DB 260 SISKEPKANWELKWLNRN-----SLMRAIWKSPFRTISVAMLYETTSDDLSSVQPPQ 312

QY 346 LKLLIKFVN--DTRAPDMQGYFYTVLLFYTACQLVLMHOYHICFVSGMRITAVIGA 402
DB 313 LRIIFDLGNBETSSKPLPLNGVFTALFLVIVSVFPLNQFYIGIEAGIGIRASL 372
QY 403 VYKALVITNSARKSSTVGEIVMLMSYDA--ORFMDLATYIMMWSAPQVILATYLLW 459
DB 373 VYKSLRLTLAENKSTGDIILMLMSYDVLRIGRFENNA--QTIIGAPQIIVLVLSV 429
QY 460 LNIQPSYLVAGVAVMTLVPNVAVMAKTKTYQVAHMSKONRIKAMEIINGIKVLKYA 519
DB 430 MLIGKAVIGGLVMTAIMMPINAFISRKVKKLSKQMKYKDMRIKTIETELLNAIKSIKYA 489
QY 520 MELAFKDKVLAIQE-ELKYLKKSAYLSANGTWTCTPPLVALCTPAVAVTIDENNILD 578
DB 490 WEEDPMARLHVHNDKELNFRKIGIVSNLIYFAMNCVPLMVCSTGFLP-SLESDPLS 548
QY 579 AQTAFFVALFNLIRPLNLTLPWVSIISVQASVSLKRLIFLSHEELPDSIER-RPVKD 637
DB 549 PAIVFPLSLPILNSAIYVSPMINITITSVSMERKLSFLSDEIDDSFIERIDPSAD 608
QY 638 GGGTNSITVNNATFTW-----ARSDPPTLNGIT-----FSIPEGALV 674
DB 609 ERALPAIEMNNITFLMKSKEVLTSQSGDVLRTDEBSIIGSSQIALKNIDHFEAKRGDLV 668
QY 675 AVYGVCCGCKSSILSALLAMDKVEGH-----VAIKSVAVYVQQAIIQNDSLRENT 726
DB 669 CVVGRVAGKSTPLKAILGOLPCMSGRDSIPKLIIRSSSVAYVCSQESVIMAVRENI 728
QY 727 LFGCOLBEPPYRSVTOACALLPDLLEIPSGDRTEIGEKVNLGGGQKORVSLARAVSNA 786
DB 729 LFGHKPDQDYDITIRACQLLPDLKILPDDELTVEKGISLGGQKARSLARAVSRA 788
QY 787 DIYLPDPLSAVDAAVAKHIFENV-IGPKMLKNKTRILVTHSMYSYLPQVDVITVNSGK 845
DB 789 DIYLLDDISAVDAEVSKNITIEVVLIGKTKALNKKTLITVTNVSLILKSKQMLYALENGE 848
QY 846 ISMGYSQIELLA--DGAFAELRTYAS-----TEQEDAEENGVTGSGPGKEAKOME 897
DB 849 IYEQGVYEDVMNKNNTSKLKLIEBFDSPIDNGESDVQTEHRSSEVDEP----- 900
QY 898 NGMLVYDSAGKOLROLOLSSSSYSQDISRHNSHTAELOKAEKKEFWKL----- 947
DB 901 -----IQLVNTESETDEVLT--ESELIELKANSRRASLATLRPPFYGAQL 945
QY 948 -----MEADKAQQTQVNLVYVMDYMKAIIGLFIISFLIFLV-CNHSVALASNYLSLW 999
DB 946 DSVKTKAQKAEKTEVGRVKTKIYLAITYIKCGV-LGVVLPFLPMLTRVPDLAENPWLKYW 1004
QY 1000 TDDPIYNGIOEHTRKRLSYGALGISQGIYAFGYSMAVSI--GGILASRCLHVDLLHSI 1056
DB 1005 SESENERKSGNERVWMEGVAYSLIGVAS--AAFNLLRSIMMLLYCSIRGSKJHESMAKV 1062
QY 1057 LRSPMSFFERTPGNIVNRPFSKELDTVDSMIPVIMFGSLFNVGACIVILLAPPIA 1116
DB 1063 IRSPMTFFETTPGRIINRPSSDMDAVDSLQYISFPFKSILFTYVTLVLYVNNPWL 1122
QY 1117 IILPGLIYFFVQRYFVASRQLKRLAESYSRSPVYSHNETLIGVSVIRAFEOERFIH 1176
DB 1123 VPMPLVLIYIYQTYIYLSRELKRLISISYSPIMSLMSSESLNGSIIDAYHFRFTY 1182
QY 1177 QSDLKVDENOKAYPSIVANRWLAFLVCEVNCIVLPALFAVI--SRHSLSAGLVGS 1233
DB 1183 LNYEKIQYVNDVFENRSTNRWLSVLOTIGATIVATATAIILATMNTKQOLSSGNGVL 1242
QY 1234 VSYSLQVTTYLMMVLVMSSEMETNIVAVERLKEYSYTBKAPMQIOETAPPSMPOGVY 1293
DB 1243 MSYSLEVTGSLTYIATTYTITINIVSVRIYVYCGLPPELAGSINERKPPDEMVPKSGCI 1302
QY 1294 EPRNYCLARREDLDPLVLRHINTINGEKVYGIYGTGAKSSITLGLPRINISABCEIIT 1353
DB 1303 EFKNYSTKYRENLDPLVLRHINTINGEKVYGIYGTGAKSSITLGLPRILBETBEKIIT 1362

QY 1354 DGINIAKIGLHDLRFKTIIPDPVLPFGSLRMNLDPPSOYSDEEWTSTLELAHKDFVS 1413
 Db 1363 DGIDISDIGFBLRSHLAIIPQDAQAFEGTVKTNLDPFNRYSEDELKRAVEQAHKPHLE 1422
 QY 1414 AL-----PDKLDHECAEGENLSVGOROLVCJARALIRKTKIIVLDEA 1456
 Db 1423 KMLHSEPRGDDSNEDGNVNDILDVKINENGSNLSVGOROLLCLARALLNRSKILVLDEA 1482
 QY 1457 TAAVLEETDLDIOSTIRTOFEDCTVLTAHRLANTIMDYTRVIVLDKGEIOEYGAPSDILQ 1516
 Db 1483 TASVDMETDKIIOPTIRREFKDRITLITAHRIDIVLDSDKIIVLDQGSVREFDPSKLLS 1542
 QY 1517 QR-GLFYSMAXDAG 1529
 Db 1543 DKTSIFYSLCEKGG 1556

Search completed: December 15, 2005, 15:36:25
 Job time : 56.1218 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:56:46 ; Search time 238.465 Seconds

(without alignments)
5594.767 Million cell updates/sec

Title: US-10-665-283-4

Perfect score: 9734

Sequence: 1 MALRGFCASDGLMDMNV.....NTIKVPTPLCTARQLDSDRS 1891

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Uniprot_05.80.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	80.7	1531	1 MRPL_HUMAN	P33527 homo sapien
2	7769	79.8	1515	2 O9UQ9_HUMAN	O9UQ9 homo sapien
3	7695	79.1	1531	1 O864R9_MACFA	O864R9 macaca fasc
4	7689	79.0	1531	2 O864S0_MACFA	O864S0 macaca fasc
5	7444	76.5	1459	2 O9UQ97_HUMAN	O9UQ97 homo sapien
6	7419.5	76.2	1456	2 O9UQAO_HUMAN	O9UQAO homo sapien
7	7272	74.7	1531	2 O6UR05_CANFA	O6UR05 canis fami
8	7242	74.4	1439	2 O59G19_HUMAN	O59G19 homo sapien
9	7230.5	74.3	1530	2 O8HX05_BOVIN	O8HX05 bos taurus
10	7104.5	73.0	1400	2 O9UQ08_HUMAN	O9UQ08 homo sapien
11	7002.5	71.9	1528	1 MRPL_MOUSE	O35379 mus musculu
12	6932.5	71.2	1532	2 O810E4_RAT	O810E4 rattus norv
13	6932.5	71.2	1532	2 O8CG09_RAT	O8CG09 rattus norv
14	6892	70.8	1523	2 O810G9_RAT	O810G9 rattus norv
15	6301	64.7	1525	2 O5F3E4_CHICK	O5F3E4 gallus gall
16	6075	62.4	1215	2 O68CP7_HUMAN	O68CP7 homo sapien
17	4526	46.5	1522	2 O5SUP4_MOUSE	O5SUP4 mus musculu
18	4522	46.5	1519	2 O80ZK8_MOUSE	O80ZK8 mus musculu
19	4522	46.5	1522	2 O56PH0_MOUSE	O56PH0 mus musculu
20	4508.5	46.3	1523	2 O59DU0_MOUSE	O59DU0 mus musculu
21	4491.5	46.1	1527	1 MRP3_HUMAN	O15438 homo sapien
22	4484.5	46.1	1533	2 O59H05_HUMAN	O59H05 homo sapien
23	4422	45.4	1514	2 O96QA9_HUMAN	O96QA9 mus musculu
24	4402	45.2	1498	2 O59DK9_MOUSE	O59DK9 mus musculu
25	4396	45.2	1522	1 MRP3_RAT	O88563 rattus norv
26	4148.5	42.6	1496	2 O4S7E7_TETNG	O4S7E7 tetraodon n
27	3885.5	39.9	1548	2 O7KTC3_DROME	O7KTC3 drosophila
28	3880.5	39.9	1548	2 O7KTC1_DROME	O7KTC1 drosophila
29	3879.5	39.8	1548	2 O7KTC2_DROME	O7KTC2 drosophila
30	3874.5	39.7	1548	2 O917N0_DROME	O917N0 drosophila
31	3860.5	39.7	1548	2 O917N0_DROME	O917N0 drosophila

32	3857	39.6	1549	2 O7KTC0_DROME	O7KTC0 drosophila
33	3845.5	39.5	1548	2 O8T9C5_DROME	O8T9C5 drosophila
34	3838.5	39.4	1548	2 O7KTB9_DROME	O7KTB9 drosophila
35	3836.5	39.4	1548	2 O7KTB9_DROME	O7KTB9 drosophila
36	3818.5	39.2	1548	2 O7KTC0_DROME	O7KTC0 drosophila
37	3813.5	39.2	1548	2 O7KTC0_DROME	O7KTC0 drosophila
38	3812.5	39.2	1548	2 O7KTC4_DROME	O7KTC4 drosophila
39	3807.5	39.1	1548	2 O7KTC9_DROME	O7KTC9 drosophila
40	3793.5	39.0	1548	2 O7KTC7_DROME	O7KTC7 drosophila
41	3790	38.9	1549	2 O9VK56_DROME	O9VK56 drosophila
42	3771.5	38.7	1548	2 O7KTC5_DROME	O7KTC5 drosophila
43	3769.5	38.7	1548	2 O7KTC6_DROME	O7KTC6 drosophila
44	3741	38.4	1564	2 O8OC98_RAJER	O8OC98 raja erinac
45	3738.5	38.4	1567	2 O6PH26_BRARE	O6PH26 brachydanio

ALIGNMENTS

RESULT 1

MRPL_HUMAN STANDARD; PRT; 1531 AA.

ID MRPL_HUMAN

AC P33527; O14819; P78419;

DT 01-FEB-1994 (Ref. 28, Created)

DT 01-NOV-1997 (Ref. 35, Last sequence update)

DT 13-SEP-2005 (Ref. 48, Last annotation update)

DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).

GN Name=ABCC1; Synonyms=MRP, MRPL;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=93088080; PubMed=1360704;

RA Cole S.P.C., Bhargava J.G., Gerlach J.H., Mackie J.E., Grant C.E., Almqvist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G., RT lung cancer cell line."

RT Science 258:1650-1654(1992).

RL [2]

RP SEQUENCE REVISION.

RX MEDLINE=93262415; PubMed=8098549;

RA Cole S.P.C., Deeley R.G.;

RT "Multidrug resistance-associated protein: sequence correction.";

RL Science 260:879-879(1993).

RL [3]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=96251691; PubMed=8649356;

RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;

RT "Structure and expression of the messenger RNA encoding the murine 'multidrug resistance protein, an ATP-binding cassette transporter.'";

RL Mol. Pharmacol. 49:962-971(1996).

RL [4]

RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.

RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;

RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;

RT "Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA.";

RL Genomics 45:368-378(1997).

RL [5]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA] OF 1131-1531.

RX MEDLINE=99452270; PubMed=10493829; DOI=10.1006/geno.1999.5927;

RA Loftus B.J., Kim U.-J., Shedd V.P., Kalush F., Brandon R., Fuhrmann U., Mason T., Crosby M.L., Barnstead M., Cronin L., Eschler B.E., Harris P.C., Venter J.C., Adams M.D.;

RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";

RL Genomics 60:295-308(1999).

[6]
RP TOPOLOGY AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=97442425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;
RA Hefner D.R., Almqvist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,
Deley R.G., Cole S.P.C.; The multidrug resistance protein (MRP). A study
of glycosylation-site mutants reveals an extracytosolic NH2
terminus.";
RL J. Biol. Chem. 272:23623-23630(1997).
[7]
RP TOPOLOGY.
RX MEDLINE=97476245; PubMed=9334225; DOI=10.1074/jbc.272.42.26479;
RA Kast C., Gros P.;
RT "Topology mapping of the amino-terminal half of multidrug resistance-
associated protein by epitope insertion and immunofluorescence.";
RL J. Biol. Chem. 272:26479-26487(1997).
[8]
RP TOPOLOGY.
RX MEDLINE=98153110; PubMed=9485377; DOI=10.1021/bi972332v;
RA Kast C., Gros P.;
RT "Epitope insertion favors a six transmembrane domain model for the
carboxy-terminal portion of the multidrug resistance-associated
protein.";
RL Biochemistry 37:2305-2313(1998).
[9]
RP MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-GLU-1455.
RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;
RA Cui L., Hou Y.-X., Rjordani J.R., Chang X.-B.;
RT "Mutations of the Walker B motif in the first nucleotide binding
domain of multidrug resistance protein MRP1 prevent conformational
maturation.";
RL Arch. Biochem. Biophys. 392:153-161(2001).
[10]
RP MUTAGENESIS OF TRP-1246.
RX MEDLINE=21238219; PubMed=11278867; DOI=10.1074/jbc.M011246200;
RA Ito K., Olsen S.L., Qiu W., Deley R.G., Cole S.P.C.;
RT "Mutation of a single conserved tryptophan in multidrug resistance
protein 1 (MRP1/ABCC1) results in loss of drug resistance and
selective loss of organic anion transport.";
RL J. Biol. Chem. 276:15616-15624(2001).
[11]
RP VARIANTS GLN-633 AND VAL-671.
RX MEDLINE=2029630; PubMed=10835642; DOI=10.1038/76102;
RA Le Saux O., Urban Z., Tschuch C., Ceiszar K., Bacchelli B.,
Quagliaro D., Pasquali-Bonchetti I., Pope F.M., Richards A., Terry S.,
Bercovich L., de Paeppe A., Boyd C.D.;
RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma
elasticum.";
RL Nat. Genet. 25:223-227(2000).
[12]
RP VARIANTS VAL-671.
RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;
RA Ringgeil F., Lehwahl M.G., Cristiano A.M., Uitto J.;
RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a
transmembrane ATP-binding cassette (ABC) transporter.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006(2000).
[13]
RP VARIANTS SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.
RX MEDLINE=21578494; PubMed=11721885; DOI=10.1007/s100380170017;
RA Conrad S., Kaufmann H.-M., Ito K., Deley R.G., Cole S.P.C.,
Schrenk D.;
RT "Identification of human multidrug resistance protein 1 (MRP1)
RT mutations and characterization of a G671V substitution.";
RL J. Hum. Genet. 46:656-663(2001).
[14]
RP VARIANTS MET-117 AND LEU-1512.
RX MEDLINE=20579883; PubMed=11139250;
DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14>3.0.CO;2-F;
RA Perdu J., Germain D.P.;
RT "Identification of novel polymorphisms in the pms and MRP1 (ABCC1)
genes at locus 16p13.1 and exclusion of both genes as responsible for
RT pseudoxanthoma elasticum.";
RL Hum. Mutat. 17:74-75(2001).

[15]
RP VARIANTS SER-433, ILE-73; GLN-723 AND GLN-1058.
RX MEDLINE=21163848; PubMed=11266082;
DOI=10.1097/00008571-200103000-00008;
RA Ito S., Ieiri I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;
RT "Polymorphism of the ABC transporter gene, MDR1, MRP1 and MRP2/cMOAT,
in healthy Japanese subjects.";
RL Pharmacogenetics 11:175-184(2001).
-1- FUNCTION: May participate directly in the active transport of
drugs into subcellular organelles or influence drug distribution
indirectly.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=8;
Comment=Additional isoforms seem to exist. Experimental
confirmation may be lacking for some isoforms;
Name=Alllexons;
CC IsoId=P33527-1; Sequence=Displayed;
CC Name=Delexon-17;
CC IsoId=P33527-2; Sequence=VSP_000037;
CC Name=Delexon-18;
CC IsoId=P33527-3; Sequence=VSP_000038;
CC Name=Delexon-30;
CC IsoId=P33527-4; Sequence=VSP_000039;
CC Name=Delexon-17-18;
CC IsoId=P33527-5; Sequence=VSP_000037, VSP_000038;
CC Name=Delexon-17-30;
CC IsoId=P33527-6; Sequence=VSP_000037, VSP_000039;
CC Name=Delexon-18-30;
CC IsoId=P33527-7; Sequence=VSP_000038, VSP_000039;
CC Name=Delexon-17-18-30;
CC IsoId=P33527-8; Sequence=VSP_000037, VSP_000038, VSP_000039;
CC -1- TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear
cells.
CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
CC -1- SIMILARITY: Contains 2 ABC transporter domains.
CC -1- DATABASIS: NME-Atlas Genet. Cytogenet. Oncol. Haematol.;
WWW="http://www.inbioigen.fr/services/chromosome/Genes/MRP106.html".
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; L05628; AAB4616.1; -; mRNA.
DR EMBL; AF022853; AAB83983.1; -; Genomic DNA.
DR EMBL; AF022824; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022826; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022827; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022828; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022829; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022831; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022838; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022839; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022840; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022848; AAB83983.1; JOINED; Genomic DNA.

DR EMBL: AF022849; AAB83983.1; JOINED; Genomic DNA.
DR EMBL: AF022850; AAB83983.1; JOINED; Genomic DNA.

Query Match 80.7%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pident. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVNTSNDFTCFQNTLVWVPCGYLMACPPEFYLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTVNTSNDFTCFQNTLVWVPCGYLMACPPEFYLYSRH 60
QY 61 DRGYIQMTPLNKTALGFLIMIVCMADLFYSFEMERSGIFLAFLVPSPTLIGITLLA 120
DB 61 DRGYIQMTPLNKTALGFLIMIVCMADLFYSFEMERSGIFLAFLVPSPTLIGITLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFWVALVCAALILRSKMTALKEDAOVDLFRDITFYVYS 180
DB 121 TFLIOLERRKGVSSGIMLTFWVALVCAALILRSKMTALKEDAOVDLFRDITFYVYS 180
QY 181 LLLIQLVLSGCSDBSPFSETHDNPCESSASFLSRITFWMTGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSGCSDBSPFSETHDNPCESSASFLSRITFWMTGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEYEAL 300
DB 241 LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEYEAL 300
QY 301 IVKSPQKEMNPISLFKVLKTFGPYFLMSFFPKAIDLMFSGPOLIKLILKPVNDTAPD 360
DB 301 IVKSPQKEMNPISLFKVLKTFGPYFLMSFFPKAIDLMFSGPOLIKLILKPVNDTAPD 360
QY 361 WQGYFTVLVFWACQTLVHOVPHICFSGMRKTAIVGAVYRKALVTNSRKSTV 420
DB 361 WQGYFTVLVFWACQTLVHOVPHICFSGMRKTAIVGAVYRKALVTNSRKSTV 420
QY 421 GEIVNLMSVDAORFMDLTYINMISAPLOVILALYLLMLNGPSVLAGAVALMVPV 480
DB 421 GEIVNLMSVDAORFMDLTYINMISAPLOVILALYLLMLNGPSVLAGAVALMVPV 480
QY 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILINGIKVLKYLAMELAFKDVLAIRQELKVLK 540
DB 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILINGIKVLKYLAMELAFKDVLAIRQELKVLK 540
QY 541 KSAVLSAVGFTVWCPELVALCTFAVYVITDENIILDAOTAFAVSLAPNITLPLNLP 600
DB 541 KSAVLSAVGFTVWCPELVALCTFAVYVITDENIILDAOTAFAVSLAPNITLPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGTSITVNNATFTMARSDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGTSITVNNATFTMARSDPT 660
QY 661 LNSITTSIPFGALVAVVGVGCCGKSLLSALLAEMDVKEGHVAKGSVAVVPQAWIQND 720
DB 661 LNSITTSIPFGALVAVVGVGCCGKSLLSALLAEMDVKEGHVAKGSVAVVPQAWIQND 720
QY 721 SLRENILFGQLEBPYRSVIOACALLPDLILPBGRTIIGEKVNLSSGQKQKRSVLAR 780
DB 721 SLRENILFGQLEBPYRSVIOACALLPDLILPBGRTIIGEKVNLSSGQKQKRSVLAR 780
QY 781 AVVSNADIVLPDPLSANDVHGKHPENYIGPKMLKNTIRILVTHSMGYLPQVDYIV 840
DB 781 AVVSNADIVLPDPLSANDVHGKHPENYIGPKMLKNTIRILVTHSMGYLPQVDYIV 840
QY 841 MSGGKISEMGSYOBELARDGAFELRTYASTQEQDAEENGVTGVSGPKKAKOMENGM 900
DB 841 MSGGKISEMGSYOBELARDGAFELRTYASTQEQDAEENGVTGVSGPKKAKOMENGM 900
QY 901 LVYDSAGKQIQRQLSSSSSYSGDISRHNSITAEIQKAEKEETWKLMEADKQGTGVKL 960
DB 901 LVYDSAGKQIQRQLSSSSSYSGDISRHNSITAEIQKAEKEETWKLMEADKQGTGVKL 960
QY 961 SYVWDYKAIQGLFISFLIFMCMNHVSALASNWYLSLMTDDPIVNGTQEHKTVRLSVYG 1020
DB 961 SYVWDYKAIQGLFISFLIFMCMNHVSALASNWYLSLMTDDPIVNGTQEHKTVRLSVYG 1020

DB 961 SYVWDYKAIQGLFISFLIFMCMNHVSALASNWYLSLMTDDPIVNGTQEHKTVRLSVYG 1020
QY 1021 ALGISQIAVFGYMAVNSIGIILASRCHADLHSLIRSPMSFEPRPSGNLVNRSKEL 1080
DB 1021 ALGISQIAVFGYMAVNSIGIILASRCHADLHSLIRSPMSFEPRPSGNLVNRSKEL 1080
QY 1081 DTVDSMPEVYIKFMGSLFVNIACIVILATPILAIITIPPLGIYFFVORFYVASSROL 1140
DB 1081 DTVDSMPEVYIKFMGSLFVNIACIVILATPILAIITIPPLGIYFFVORFYVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQAXYPSIVANWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQAXYPSIVANWLA 1200
QY 1201 VRLCVCNCIVLPALFVAVISRLSLAGLVGSYSLOVTTIYNTLVNRSSEMETNIVA 1260
DB 1201 VRLCVCNCIVLPALFVAVISRLSLAGLVGSYSLOVTTIYNTLVNRSSEMETNIVA 1260
QY 1261 VERLKEYSETKEKAPMOIOETAPSSWPQYGRVFRNYCLRYREDLDFVLRIHNTVING 1320
DB 1261 VERLKEYSETKEKAPMOIOETAPSSWPQYGRVFRNYCLRYREDLDFVLRIHNTVING 1320
QY 1321 EKVIGVGTGAGKSSLTGLFRINESABGEIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVIGVGTGAGKSSLTGLFRINESABGEIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRNMLDPPSOYSDEEVTSLSLAHLKDFVSNLPKLDHECAEGENLSVGQROLVCL 1440
DB 1381 SGSLRNMLDPPSOYSDEEVTSLSLAHLKDFVSNLPKLDHECAEGENLSVGQROLVCL 1440
QY 1441 ARALLRKTKLTVDEATAADLTDDLOSTIRTOFEDCTVLTAHNLNTIMOTRYIVL 1500
DB 1441 ARALLRKTKLTVDEATAADLTDDLOSTIRTOFEDCTVLTAHNLNTIMOTRYIVL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYGMADAGIV 1531
DB 1501 DKGEIOEYGA PSDLLQORGLFYGMADAGIV 1531

RESULT 2
Q9U099 HUMAN
ID Q9U099 HUMAN PRELIMINARY; PRT: 1515 AA.
AC Q9U099
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA."
RL Genomics 45:368-378 (1997).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AF022853; AAB83979.1; -; Genomic DNA.
DR EMBL: AF022827; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022828; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022829; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022831; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022833; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022835; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022837; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022843; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022845; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022849; AAB83979.1; JOINED; Genomic DNA.

DR EMBL; AF022848; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022852; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022851; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022840; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022838; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022826; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022824; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83979.1; JOINED; Genomic DNA.
DR HSP; P08716; IMTO.
DR Ensembl; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP assoc.
DR InterPro; IPR00719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR Prodom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA_2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMlf; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding.
FT NON_TER 1 1
SQ SEQUENCE 1515 AA; 169853 MW; DC85592817C439FE CRC64;

Query March 79.8%; Score 7769; DB 2; Length 1515;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DMNTVNTSNPDPFKCFQNTVLVWPCFYLMACFPFYFLYSRHDGTYIQMTPLNKTXTA 76
DB 1 DMNTVNTSNPDPFKCFQNTVLVWPCFYLMACFPFYFLYSRHDGTYIQMTPLNKTXTA 60
QY 77 LGFLIMIVCWADLFYFSEWERSRGIFLAPVLVSPFLIGITTLATFLLQLERRKGVSSG 136
DB 61 LGFLIMIVCWADLFYFSEWERSRGIFLAPVLVSPFLIGITTLATFLLQLERRKGVSSG 120
QY 137 IMTLFWVAIVCALIIRSKTMTALKEDAVDLPDITFFVYFSLLLQLVLSCFSDSP 196
DB 121 IMTLFWVAIVCALIIRSKTMTALKEDAVDLPDITFFVYFSLLLQLVLSCFSDSP 180
QY 197 LFEETIHDPNCPRESSASFLSRITFWMTTGLVNGYROPFLSGSDLSLNKEDTSQVVPV 256
DB 181 LFEETIHDPNCPRESSASFLSRITFWMTTGLVNGYROPFLSGSDLSLNKEDTSQVVPV 240
QY 257 LVKWKKECAKTRKQPVYVSSKDPAPQKSSKVDANEVEALIVKSPQKEMPSPFKV 316
DB 241 LVKWKKECAKTRKQPVYVSSKDPAPQKSSKVDANEVEALIVKSPQKEMPSPFKV 300
QY 317 LYKTFGYPFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPDMOGIFTYVLLFVTAQL 376
DB 301 LYKTFGYPFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPDMOGIFTYVLLFVTAQL 360

QY 377 QTVLHGYFHCFSVGMRIKTAIVIGAVYRKALVYITNSARKSVYGEIVNLMSVDQRPM 436
DB 361 QTVLHGYFHCFSVGMRIKTAIVIGAVYRKALVYITNSARKSVYGEIVNLMSVDQRPM 420
QY 437 LATYINNIWASAPQVILATVLLMLNIGPSVLAQVAVVWLVPPVNAVAMTKTKYQVAHMK 496
DB 421 LATYINNIWASAPQVILATVLLMLNIGPSVLAQVAVVWLVPPVNAVAMTKTKYQVAHMK 480
QY 497 SKDNRIKLAMEIINGIKVLKLYAMELAFKDKVLAIROEELKVLKKSAYLSAVGTFTVCT 556
DB 481 SKDNRIKLAMEIINGIKVLKLYAMELAFKDKVLAIROEELKVLKKSAYLSAVGTFTVCT 540
QY 557 PFVVALCTPAVYVITDENNLIDQTAFAVSIALFNILRFPNLIPMTVSSIVQASVSLKRL 616
DB 541 PFVVALCTPAVYVITDENNLIDQTAFAVSIALFNILRFPNLIPMTVSSIVQASVSLKRL 600
QY 617 RIFLSHEELPPDSTERPVVDGGTNSITVRNATFTWARSDDPTLNGITFSIPGALVAV 676
DB 601 RIFLSHEELPPDSTERPVVDGGTNSITVRNATFTWARSDDPTLNGITFSIPGALVAV 660
QY 677 VGVGCGKSSLLSALLAEMDKVEGVNAIKGSVAVVPQAMIQNDSLRENLIFGQLEBPY 736
DB 661 VGVGCGKSSLLSALLAEMDKVEGVNAIKGSVAVVPQAMIQNDSLRENLIFGQLEBPY 720
QY 737 YRSVIAQCALLPDLEILPSGDRTEIGKGVNLSCGQKQSVSLARAVYSNADIVLPDPLS 796
DB 721 YRSVIAQCALLPDLEILPSGDRTEIGKGVNLSCGQKQSVSLARAVYSNADIVLPDPLS 780
QY 797 AYDAHNGKHIFENVIGPKMLKXKTRILYTHSMSTYPOUVIITYMSGGKISEMSYQEL 856
DB 781 AYDAHNGKHIFENVIGPKMLKXKTRILYTHSMSTYPOUVIITYMSGGKISEMSYQEL 840
QY 857 ARDGAFAEFLRTYASTEQEODAEENGTVGSGPGEAKQKQENGLVYDSAGKQLORLSS 916
DB 841 ARDGAFAEFLRTYASTEQEODAEENGTVGSGPGEAKQKQENGLVYDSAGKQLORLSS 900
QY 917 SSSYSGDISRRHNSTAELOAKAEKEETWKLMEADKAQTCQVKLSYWDYWKAIGLFISF 976
DB 901 SSSYSGDISRRHNSTAELOAKAEKEETWKLMEADKAQTCQVKLSYWDYWKAIGLFISF 960
QY 977 LSTLFPMQNVSAASAANYWMLSLMTDDPIVNGCTQHTVRSLSYVCAIGISQIAIFGVSMA 1036
DB 961 LSTLFPMQNVSAASAANYWMLSLMTDDPIVNGCTQHTVRSLSYVCAIGISQIAIFGVSMA 1020
QY 1037 VSIGGILASRCLAHVDLHSLIRSPMSFFERTPSGNLVNRSKELDYDMSIPEVIKAFMG 1096
DB 1021 VSIGGILASRCLAHVDLHSLIRSPMSFFERTPSGNLVNRSKELDYDMSIPEVIKAFMG 1080
QY 1097 SLEFNIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQIKRLSVSRSPVSHFN 1156
DB 1081 SLEFNIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQIKRLSVSRSPVSHFN 1140
QY 1157 ETLIGSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRMALVRLCVCNGCIVLPAAL 1216
DB 1141 ETLIGSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRMALVRLCVCNGCIVLPAAL 1200
QY 1217 PAVISRSLSAGVLGVSYSLOVTTYLNMVLRMSSEMETIYAVERLKEYSSETEKAPW 1276
DB 1201 PAVISRSLSAGVLGVSYSLOVTTYLNMVLRMSSEMETIYAVERLKEYSSETEKAPW 1260
QY 1277 QIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHNTVINGEKVGI VGRTAGKSSL 1336
DB 1261 QIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHNTVINGEKVGI VGRTAGKSSL 1320
QY 1337 TLGLFRINESAGEIITIDGINIAKIGLHDLRFKTTIIPDPVLFSGSIRNMLDPFSQYSD 1396
DB 1321 TLGLFRINESAGEIITIDGINIAKIGLHDLRFKTTIIPDPVLFSGSIRNMLDPFSQYSD 1380
QY 1397 BEVWTSLEIAHLKQFVSLPDKLDHECABEGENLSVQORQVLCARALLRTKTLIVLDEA 1456
DB 1381 BEVWTSLEIAHLKQFVSLPDKLDHECABEGENLSVQORQVLCARALLRTKTLIVLDEA 1440
QY 1457 TAAVDLETDLLIOSTIRTOPEDCVTLLTAHRLANTIMDYTRVIVLDKGEIOETGAPSDLLQ 1516

Db 1441 TAAVLETDLDLQSTIRQFEDCTVLTIAHRLNTIMDTYRIVLDKGEIQGYGAPSDLLQ 1500
Qy 1517 ORGLFYSMADAGLV 1531
Db 1501 ORGLFYSMADAGLV 1515

RESULT 3
08649 MACFA
ID 08649 MACFA PRELIMINARY; PRT; 1531 AA.
AC 08649;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Multidrug resistance protein 1B.
OS Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22544876; PubMed=12657726;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL M01. Cancer Ther. 2:307-316(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Perry W.L., III, Godinot N.;
RL Submitted (Aug-2002) to the EMBL/GenBank/DDBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY146673; AAN65349.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1531 AA; 171659 MW; 1AE78BEFDF9EF459 CRC64;

Query Match 79.1%; Score 7695; DB 2; Length 1531;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

Qy 1 MALRGFCADGSDPLMDMNTVNTSNDFTKCFQNTVLVWVPCFYLMACEPFYFLYSRH 60
Db 1 MALRGFCADGSDPLMDMNTVNTSNDFTKCFQNTVLVWVPCFYLMACEPFYFLYSRH 60
Qy 61 DRGVIQMTPLNKTITAGFLIMTYCAGDLFYSFWEKRGIFLAVPLVSPITLIGITTLA 120
Db 61 DRGVIQMTPLNKTITAGFLIMTYCAGDLFYSFWEKRGIFLAVPLVSPITLIGITTLA 120

Qy 121 TFLIQERRKGVSSGIMLFWMVAIVCALAIILRSKITMTALKEDAOVDFRDIPTFYYS 180
Db 121 TFLIQERRKGVSSGIMLFWMVAIVCALAIILRSKITMTALKEDAOVDFRDIPTFYYS 180
Qy 181 LLIQLVLSGFCSDRSPFLFSETIHDNPNCPSSASFLSRITFWMTGLIVRGYRPLEGSD 240
Db 181 LVIQLVLSGFCSDRSPFLFSETIHDNPNCPSSASFLSRITFWMTGLIVRGYRPLEGSD 240
Qy 241 LMSLNKEDTSBOVVPVTVKMKKECATRQKQPVVYVSSDPAQPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSBOVVPVTVKMKKECATRQKQPVVYVSSDPAQPKSSKVDANEVEAL 300
Qy 301 IVSPQKEMNPSPFKVLYKTFGPFYFLMSFFFKAIHDMFSGPOILKLLIKFVNDTAPD 360
Db 301 IVSPQKEMNPSPFKVLYKTFGPFYFLMSFFFKAIHDMFSGPOILKLLIKFVNDTAPD 360
Qy 361 MGGFYTVLLFVTAQTLVLHQYFHIQFVSGMRKTAVIGAVYRKALVITNSARKSTV 420
Db 361 MGGFYTVLLFVTAQTLVLHQYFHIQFVSGMRKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLSVDAQRPMDLATTINMWSAPLOYIALYLLMLNLPVSLAGVAVMLMPVN 480
Db 421 GEIVNLSVDAQRPMDLATTINMWSAPLOYIALYLLMLNLPVSLAGVAVMLMPVN 480
Qy 481 AVNAMKTKTVQVAHMKSKDKRILKMEILNGIKVLKLYAMEIAFKDVLAIROBELKYLK 540
Db 481 AVNAMKTKTVQVAHMKSKDKRILKMEILNGIKVLKLYAMEIAFKDVLAIROBELKYLK 540
Qy 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVITIDENNILDAOTAFAVSLALFNILRPILNP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVITIDENNILDAOTAFAVSLALFNILRPILNP 600
Qy 601 MVTSSIQASVSLKRLRILFISHBLEPDSIERRPVQDGGTNSITVENAFTMARSDPT 660
Db 601 MVTSSIQASVSLKRLRILFISHBLEPDSIERRPVQDGGTNSITVENAFTMARSDPT 660
Qy 661 LNGITTSIPRGALVAVVGVGCGKSLSLALBMDVBEHVAKSGVAVPOQAMQND 720
Db 661 LNGITTSIPRGALVAVVGVGCGKSLSLALBMDVBEHVAKSGVAVPOQAMQND 720
Qy 721 SLRENILFGCLLEPPYRSYIOACALPDEILIPSGDRTIEGKVMISGQOKORVSLAR 780
Db 721 SLRENILFGCLLEPPYRSYIOACALPDEILIPSGDRTIEGKVMISGQOKORVSLAR 780
Qy 781 AVYSNADTYLPDPLSANDAHVGHIFENYIGPKMKKNTRIILVTHSMGYLPQVDVIY 840
Db 781 AVYSNADTYLPDPLSANDAHVGHIFENYIGPKMKKNTRIILVTHSMGYLPQVDVIY 840
Qy 841 MSGGKISEMGSYQELLARDGAFAPLRTVASTQEOQDAEENGVTGVSQPKKAKOMENG 900
Db 841 MSGGKISEMGSYQELLARDGAFAPLRTVASTQEOQDAEENGVTGVSQPKKAKOMENG 900
Qy 901 LVYDSAGKQLORQLSSSSSYSGDISRHNNSTAELOKAEAKKETWKLMEADKAQOGVKL 960
Db 901 LVYDSAGKQLORQLSSSSSYSGDISRHNNSTAELOKAEAKKETWKLMEADKAQOGVKL 960
Qy 961 SVYWDYKAKIAGLIFSILSIFLPMCNHVSALASNYWLSLWTDPIVNGTQHTYRSLVYG 1020
Db 961 SVYWDYKAKIAGLIFSILSIFLPMCNHVSALASNYWLSLWTDPIVNGTQHTYRSLVYG 1020
Qy 1021 ALGISGCIANRGYSMAVSTIGLILASRCIHLVDLHSLIRSMSPFERPSGNLVNRSKEL 1080
Db 1021 ALGISGCIANRGYSMAVSTIGLILASRCIHLVDLHSLIRSMSPFERPSGNLVNRSKEL 1080
Qy 1081 DTYDSMIPVYIKMFMSLFFVIGACIYLLATPILAIIPPLGLIYFFVORFYAASROL 1140
Db 1081 DTYDSMIPVYIKMFMSLFFVIGACIYLLATPILAIIPPLGLIYFFVORFYAASROL 1140
Qy 1141 KRLSVSRSPVYSHFNMTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYYPSIVANRWLA 1200
Db 1141 KRLSVSRSPVYSHFNMTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYYPSIVANRWLA 1200
Qy 1201 VRLQCVNCIVLPAALFAVYSRSLSAGVLGVSYSLGVTYTLNMLVRRSMSEMETIIVA 1260

Db 1201 VRLECVNCIVLFAALFRAVLSRHSLSAGLVGLSVSYSLQVTTYTNMVLKMSMEETIVA 1260
QY 1261 VERLKEVSETEKEAPWQIOETAPSPSWPQVGRVEFRNCLARYEDLDVFLHINVTNGG 1320
Db 1261 VERLKEVSETEKEAPWQIOETAPSPSWPQVGRVEFRNCLARYEDLDVFLHINVTNGG 1320
QY 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIGDINIAKIGLDLRPKITIIPODPVL 1380
Db 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIGDINIAKIGLDLRPKITIIPODPVL 1380
QY 1381 SGTIRNMLDPPRSQVSDERWMTSLFLAHKDFVSALPDLHHECAAGGNLSVGQOLVCL 1440
Db 1381 SGTIRNMLDPPRSQVSDERWMTSLFLAHKDFVSALPDLHHECAAGGNLSVGQOLVCL 1440
QY 1441 ARALLRKTKIIVLEATPAVVDLETDLDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
Db 1441 ARALLRKTKIIVLEATPAVVDLETDLDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIOEYGAPSDILLQORGLFYMAKDAVL 1531
Db 1501 DKGEIOEYGAPSDILLQORGLFYMAKDAVL 1531

RESULT 4

Q864S0_MACFA PRELIMINARY; PRT; 1531 AA.
ID Q864S0;
AC Q864S0;
DT 01-JUN-2003 (TREMblrel. 24, Created)
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Multidrug resistance protein 1A.
GN Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecidae; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22544876; PubMed=12657726;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Perry W.L., Ili, Godinot N.;
RL Submitted (AUG-2002) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the MRP transporter family.
DR EMBL; AY146672; AAN5348.1; -; mRNA.
DR HSSP; P08716; IMT0.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.

DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS0107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1531 AA; 171701 MW; 168712B8CC2DB89 CRC64;

Query Match 79.0%; Score 7689; DB 2; Length 1531;
Best local similarity 97.6%; Pred. No. 0;
Matches 1494; Conservative 16; Mismatches 19; Indels 0; Gaps 0;

QY 1 MALRGFCSAGSDPLMDMNTWTNTSNPDFTKCFONTLVWVPCGYLWACPFETLYLSRH 60
Db 1 MALRGFCSAGSDPLMDMNTWTNTSNPDFTKCFONTLVWVPCGYLWACPFETLYLSRH 60
QY 61 DRGIQMTPLNKTKTALGFLLMVICMADLFYSFWERSRGIFLAVFLVSPITLIGITLLA 120
Db 61 DRGIQMTPLNKTKTALGFLLMVICMADLFYSFWERSRGIFLAVFLVSPITLIGITLLA 120
QY 121 TFLIQERRKGVQSSGIMLTFWLVVALCALILRSKIMTALKEDAVDLEFDITFYFYS 180
Db 121 TFLIQERRKGVQSSGIMLTFWLVVALCALILRSKIMTALKEDAVDLEFDITFYFYS 180
QY 181 LLLIQVLSCFSDSPFLSETIHDNPPCPSSASFLSRITFWMTTGILVGRQPLSGSD 240
Db 181 LVLIQVLSCFSDSPFLSETIHDNPPCPSSASFLSRITFWMTTGILVGRQPLSGSD 240
QY 241 LMSLNKEDTSEQVVPVLVVKWKECAKTRKQPVKVYSSKDPAPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVVKWKECAKTRKQPVKVYSSKDPAPKSSKVDANEVEAL 300
QY 301 IVKSPQKBNDSLFKVIYKTFGPYFLMSFFPKAIHDIIMMSGPOILKLIKFNVDYKAPD 360
Db 301 IVKSPQKBNDSLFKVIYKTFGPYFLMSFFPKAIHDIIMMSGPOILKLIKFNVDYKAPD 360
QY 361 WQGFYFVLLFVTPACLOTIVLHOYFHI CFSGMIKTAIVGAVYRKALVTNSARKSTV 420
Db 361 WQGFYFVLLFVTPACLOTIVLHOYFHI CFSGMIKTAIVGAVYRKALVTNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMISAPLQVITLALYLLMLAGPSVLAVAVMYLMPVN 480
Db 421 GEIVNLSVDAQRFMDLATYINMISAPLQVITLALYLLMLAGPSVLAVAVMYLMPVN 480
QY 481 AVMAKTKTYQAVAHMKSNDNRKILMNEILINGIKVLKLYAMELAFKDYALAIROBELKYLK 540
Db 481 AVMAKTKTYQAVAHMKSNDNRKILMNEILINGIKVLKLYAMELAFKDYALAIROBELKYLK 540
QY 541 KSAVLASVGFPTWCCTPEVALCTFAVYVTDENNIIDAQAFSLALFNLRPEPLNLP 600
Db 541 KSAVLASVGFPTWCCTPEVALCTFAVYVTDENNIIDAQAFSLALFNLRPEPLNLP 600
QY 601 MVISSIVQASVSLKRLIFLSHELEPDSIERRPVKQGGTNSITVRNATPTMARSDPT 660
Db 601 MVISSIVQASVSLKRLIFLSHELEPDSIERRPVKQGGTNSITVRNATPTMARSDPT 660
QY 661 LINGITFSPREGALVAVVGVQVCGKSSLLSALLAEMDRVEGHVAKGSVAVVYPOQAMIOND 720
Db 661 LINGITFSPREGALVAVVGVQVCGKSSLLSALLAEMDRVEGHVAKGSVAVVYPOQAMIOND 720
QY 721 SLRENILFGCLLEPPYRSVYQACALPLDLEILSGRTEIGEKNVLSGGQKRVSLAR 780
Db 721 SLRENILFGCLLEPPYRSVYQACALPLDLEILSGRTEIGEKNVLSGGQKRVSLAR 780
QY 781 AVYENADIVLEPDDPLASVAVDAVGHIFENVIGPGMLKNKRIIVTSHMSYLPQVDVTV 840
Db 781 AVYENADIVLEPDDPLASVAVDAVGHIFENVIGPGMLKNKRIIVTSHMSYLPQVDVTV 840
QY 841 MSGKISEMSGYQELARDAFAFLRTYASTBOEODAEENGVTGVSGPGKEAKOMENG 900
Db 841 MSGKISEMSGYQELARDAFAFLRTYASTBOEODAEENGVTGVSGPGKEAKOMENG 900
QY 901 LVYDSAGKOLOROLSSSSSYSGDLSRHNSYTAELQKAKKEETWKLMEADKATQGVKL 960
Db 901 LVYDSAGKOLOROLSSSSSYSGDLSRHNSYTAELQKAKKEETWKLMEADKATQGVKL 960

QY 961 SYVDYKAKIQLFISPLSIFLFCMCHVASALSNYWLSLMTDDPIVNGTQHEHTKRLSYG 1020
 DB 961 SYVDYKAKIQLFISPLSIFLFCMCHVASALSNYWLSLMTDDPIVNGTQHEHTKRLSYG 1020
 QY 1021 ALGISQGIAGVGSMAVSIQGIILASRCLAVDLHLHSILSPSPFFERPSGNLVNRSKEL 1080
 DB 1021 ALGISQGIAGVGSMAVSIQGIILASRCLAVDLHLHSILSPSPFFERPSGNLVNRSKEL 1080
 QY 1081 DTYSMTPEVTKMNGSLFNVICACIYLLATPTAAIIPPLGLTYFFVQPFYVASSROL 1140
 DB 1081 DTYSMTPEVTKMNGSLFNVICACIYLLATPTAAIIPPLGLTYFFVQPFYVASSROL 1140
 QY 1141 KRLSVSRSPYSHFNFTLGLSVYIRAFEEQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
 DB 1141 KRLSVSRSPYSHFNFTLGLSVYIRAFEEQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
 QY 1201 VRLFCVNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYLVNMLVRMSSEMETNIVA 1260
 DB 1201 VRLFCVNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYLVNMLVRMSSEMETNIVA 1260
 QY 1261 VERLKEVSETEKEAPMOIQETAPPSWPQVGRVPEFRNYCLRYREDLDFVLRIHVVTTNGG 1320
 DB 1261 VERLKEVSETEKEAPMOIQETAPPSWPQVGRVPEFRNYCLRYREDLDFVLRIHVVTTNGG 1320
 QY 1321 EKVGIVERTGAGKSLTGLFRINESAGEIINDGINAKTGLHDLFKITTIPODPVLF 1380
 DB 1321 EKVGIVERTGAGKSLTGLFRINESAGEIINDGINAKTGLHDLFKITTIPODPVLF 1380
 QY 1381 SGSLRMLNIDPFSSQYSDSEWMTSLBLAKDFVSALPKLDHECAEGENLSVQGRQVCL 1440
 DB 1381 SGSLRMLNIDPFSSQYSDSEWMTSLBLAKDFVSALPKLDHECAEGENLSVQGRQVCL 1440
 QY 1441 ARALLRTKILVLDDEATPAVDLETDLIQSTIRFQFEDCIVLTAHRLNTIMDYTRYIVL 1500
 DB 1441 ARALLRTKILVLDDEATPAVDLETDLIQSTIRFQFEDCIVLTAHRLNTIMDYTRYIVL 1500
 QY 1501 DKGEIOEYGA PSDLLOQRGLFYSMAXKAGLV 1531
 DB 1501 DKGEIOEYGA PSDLLOQRGLFYSMAXKAGLV 1531

DR EMBL/ AF022838; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022848; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022847; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022846; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022845; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022844; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022843; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022842; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022841; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022839; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022835; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022837; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022835; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022833; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022831; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022829; AAB83980.1; JOINED; Genomic DNA.
 DR EMBL/ AF022827; AAB83980.1; JOINED; Genomic DNA.
 DR HSBP; P08716; IMT0.
 DR EMBL; ENSG00000103222; Homo sapiens.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016887; F:ATPase activity; IEA.
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR011527; ABC membrane 1.
 DR InterPro; IPR001140; ABC_TM_transp.
 DR InterPro; IPR003438; ABC_transp_like.
 DR InterPro; IPR005292; MRP_assoc.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00664; ABC membrane; 2.
 DR Pfam; PF00005; ABC_tran; 1.
 DR Prodom; PD000006; ABC transporter; 1.
 DR SMART; SM00382; AAA; 2.
 DR TIGRfam; TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE; PS00929; ABC_TMIF; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR ATP-binding; Nucleotide-binding.
 DR NON_TER 1
 FT SEQUENCE 1459 AA; 163831 MW; 6A016A5A6E7A95C CRC64;
 Query Match 76.5%; Score 7444; DB 2; Length 1459;
 Best Local Similarity 96.3%; Pred. No. 0; Mismatches 0; Indels 56; Gaps 1;
 Matches 1459; Conservative 0; Mismatches 0; Indels 56; Gaps 1;
 QY 17 DNVVTWNTSNPDFTKCFONTLVVWPCFYLMACFPYFLYLSRHDRGYIOMTPLNKTKTA 76
 DB 1 DNVVTWNTSNPDFTKCFONTLVVWPCFYLMACFPYFLYLSRHDRGYIOMTPLNKTKTA 76
 QY 77 LGFLIMIVCAADLFYSFWEBSRGIFLAPVPLVSPPTLIGITTLATPLIQLERRKGVQSSG 136
 DB 77 LGFLIMIVCAADLFYSFWEBSRGIFLAPVPLVSPPTLIGITTLATPLIQLERRKGVQSSG 136
 QY 61 LGFLIMIVCAADLFYSFWEBSRGIFLAPVPLVSPPTLIGITTLATPLIQLERRKGVQSSG 120
 DB 61 LGFLIMIVCAADLFYSFWEBSRGIFLAPVPLVSPPTLIGITTLATPLIQLERRKGVQSSG 120
 QY 137 IMLTFLVVLVLCALILRSKIMFALKEDAQVDLFRDITTFVYVSSLILQVLVSCFSDRSP 196
 DB 137 IMLTFLVVLVLCALILRSKIMFALKEDAQVDLFRDITTFVYVSSLILQVLVSCFSDRSP 196
 QY 121 IMLTFLVVLVLCALILRSKIMFALKEDAQVDLFRDITTFVYVSSLILQVLVSCFSDRSP 180
 DB 121 IMLTFLVVLVLCALILRSKIMFALKEDAQVDLFRDITTFVYVSSLILQVLVSCFSDRSP 180
 QY 197 LPSFTIHDPNPCESSASFLSRITFWITGLIVRGYROPLEGSDLSLNKEDTSEQVVPV 256
 DB 197 LPSFTIHDPNPCESSASFLSRITFWITGLIVRGYROPLEGSDLSLNKEDTSEQVVPV 256
 QY 181 LPSFTIHDPNPCESSASFLSRITFWITGLIVRGYROPLEGSDLSLNKEDTSEQVVPV 240
 DB 181 LPSFTIHDPNPCESSASFLSRITFWITGLIVRGYROPLEGSDLSLNKEDTSEQVVPV 240
 QY 257 LVNWKKECKAKTKRKQPKVYVSSKDPAPQPKESKUNANBEVELLYKSPQKKNPSSLFKY 316
 DB 257 LVNWKKECKAKTKRKQPKVYVSSKDPAPQPKESKUNANBEVELLYKSPQKKNPSSLFKY 316
 QY 241 LVNWKKECKAKTKRKQPKVYVSSKDPAPQPKESKUNANBEVELLYKSPQKKNPSSLFKY 300
 DB 241 LVNWKKECKAKTKRKQPKVYVSSKDPAPQPKESKUNANBEVELLYKSPQKKNPSSLFKY 300
 QY 317 LVYTFGPYFLMSPFFAIIHDLMMFSGPOLIKLIIKIVNDTKAPDMQGYFTYVTLFTYACL 376
 DB 317 LVYTFGPYFLMSPFFAIIHDLMMFSGPOLIKLIIKIVNDTKAPDMQGYFTYVTLFTYACL 376

Db 301 LYKTFGEYFLMSFEFKAIHIDLMFSGPOILKLLIKFVNDTKAPDMQGYFTVILLFVTA CL 360
Qy 377 QTVLHGYFHICFSGNRIKTAIVIGAVYRKALVITNSARKSSTVGEIYNLMSVDAQRFMD 436
Db 361 QTVLHGYFHHCFSGNRIKTAIVIGAVYRKALVITNSARKSSTVGEIYNLMSVDAQRFMD 420
Qy 437 LATYNNIWSAPLQVILLALYLLMLNGPSVLGAVAVMLAVPVNAVAMAKTKTYQVAHMK 496
Db 421 LATYNNIWSAPLQVILLALYLLMLNGPSVLGAVAVMLAVPVNAVAMAKTKTYQVAHMK 480
Qy 497 SKONRIKLAMEILLNGIYVKLYAMELAPKDKVLAIRQELKVKKKSAVLSAVGFTWVCT 556
Db 481 SKONRIKLAMEILLNGIYVKLYAMELAPKDKVLAIRQELKVKKKSAVLSAVGFTWVCT 540
Qy 557 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSLAPFLIRPPLNILPMVSISSIVQASVSLKRL 616
Db 541 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSLAPFLIRPPLNILPMVSISSIVQASVSLKRL 600
Qy 617 RIFLSHEELEBDSIERPPVKDGGGTSITVRNATFTWARSDPPTLNGITFSIPGALVAV 676
Db 601 RIFLSHEELEBDSIERPPVKDGGGTSITVRNATFTWARSDPPTLNGITFSIPGALVAV 660
Qy 677 VGVQGCCSSILSLALAEKMDKVEGHVAIKGSVAVYPOQAMQNDSLRENILIFGQLEBEPY 736
Db 661 VGVQGCCSSILSLALAEKMDKVEGHVAIKGSVAVYPOQAMQNDSLRENILIFGQLEBEPY 720
Qy 737 YRSYIQAACALPDLIELPSGDRTEIGEXGVMLSGGQKRVSLARAVYSNADIYLFDDPLS 796
Db 721 YRSYIQAACALPDLIELPSGDRTEIGEXGVMLSGGQKRVSLARAVYSNADIYLFDDPLS 747
Qy 797 AVDAHVGKHLFENVIGPKMKLKNKTRILLVTHSMSTLPQVDVIIVMSGKISEMSGYOELL 856
Db 748 -----KTRILVTHSMSTLPQVDVIIVMSGKISEMSGYOELL 784
Qy 857 ARDAPAPAEPLRTYASTROBDAFENGVTGSGPGKEAQMNGMLYVTHSAAKQIQORUSS 916
Db 785 ARDAPAPAEPLRTYASTROBDAFENGVTGSGPGKEAQMNGMLYVTHSAAKQIQORUSS 844
Qy 917 SSSYSGDISRHHNSTAELQAKAEKKEETWKLEADKAOTGVKLSVYDYKAKIGLFTISF 976
Db 845 SSSYSGDISRHHNSTAELQAKAEKKEETWKLEADKAOTGVKLSVYDYKAKIGLFTISF 904
Qy 977 LSLFLFMCNHNVSALASNYWLSLMTDDPIVNGTOEHTKVRLSVYGAIGISQGIYAVFGYSMA 1036
Db 905 LSLFLFMCNHNVSALASNYWLSLMTDDPIVNGTOEHTKVRLSVYGAIGISQGIYAVFGYSMA 964
Qy 1037 VSIIGIILASRCLAYDLHLSIIRSPMSFPERTPSGNLVNRFSKELDTVDMSMPEVYIKPMFG 1096
Db 965 VSIIGIILASRCLAYDLHLSIIRSPMSFPERTPSGNLVNRFSKELDTVDMSMPEVYIKPMFG 1024
Qy 1097 SLFNVIGACIYIILATPIAAIIPPLIGLYFFVORFFVYASSRQKLKRLSVSRSPYSHFN 1156
Db 1025 SLFNVIGACIYIILATPIAAIIPPLIGLYFFVORFFVYASSRQKLKRLSVSRSPYSHFN 1084
Qy 1157 ETLGVSIVIRAFEBEOERFIHOSDKVDENOKAYPSIYANWMLVRELCVNGCIYLFAL 1216
Db 1085 ETLGVSIVIRAFEBEOERFIHOSDKVDENOKAYPSIYANWMLVRELCVNGCIYLFAL 1144
Qy 1217 PAVISRSLSLAKVGLSVSYSLQVTTYANMLVRNMSSEMETNIYVAVERLKEYSETEKEAPM 1276
Db 1145 PAVISRSLSLAKVGLSVSYSLQVTTYANMLVRNMSSEMETNIYVAVERLKEYSETEKEAPM 1204
Qy 1277 QIOETAPSSMPOVGRVFRNYCLARYEDLPVLRHIVNTLNGCKVGIIVRTGAGKSSL 1336
Db 1205 QIOETAPSSMPOVGRVFRNYCLARYEDLPVLRHIVNTLNGCKVGIIVRTGAGKSSL 1264
Qy 1337 TLGAFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMLNDPFSQYSD 1396
Db 1265 TLGAFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMLNDPFSQYSD 1324
Qy 1397 EEVWTSLELAKDFVSLPDKLDHECAEGEENISVSGRQVLCTARALLRKTIVLDEA 1456
Db 1325 EEVWTSLELAKDFVSLPDKLDHECAEGEENISVSGRQVLCTARALLRKTIVLDEA 1384

Qy 1457 TAAVDETDLLIOSTIRTOEDCTVLTIAHRLNTIMDYREVIYLDKGEIOEGAPSDLLQ 1516
Db 1385 TAAVDETDLLIOSTIRTOEDCTVLTIAHRLNTIMDYREVIYLDKGEIOEGAPSDLLQ 1444
Qy 1517 ORGLFYSMAQDAGLV 1531
Db 1445 ORGLFYSMAQDAGLV 1459
RESULT 6
ID Q9UQAO_HUMAN PRELIMINARY; PRT; 1456 AA.
AC Q9UQAO;
DC 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
CX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF022853; AAB83981.1; -; Genomic DNA.
DR EMBL; AF022852; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022826; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022827; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022829; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022831; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022840; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022838; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022852; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022851; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022849; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022848; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022828; AAB83981.1; JOINED; Genomic DNA.
DR HSSP; P08716; 1MT0.
DR Ensembl; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transport.
DR InterPro; IPR003439; ABC_transport_like.

DR InterPro; IPR005292; MRP assoc.
 DR InterPro; IPR00719; Prot_kinase.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD00006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE; PS50929; ABC_TM1f; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR ATP-binding; Nucleotide-binding.
 FT NON_TER 1
 SQ SEQUENCE 1456 AA; 163233 MW; 8DB8AB2BC481F2 CRC64;

Query Match 76.2%; Score 7419.5; DB 2; Length 1456;
 Best Local Similarity 96.1%; Pred. No. 0;
 Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;

QY 17 DMNVTMTSNDPFTKCFONTVLVWVPCFYLMACFPFYLYSRHRCGYIQMTPIKTKTA 76
 DB 1 DMNVTMTSNDPFTKCFONTVLVWVPCFYLMACFPFYLYSRHRCGYIQMTPIKTKTA 60

QY 77 LGFLMIIVCWAADLFYSFWERSRGIFLAPVFLVSPTLIGITTLATFLIOLERRKGVSSG 136
 DB 61 LGFLMIIVCWAADLFYSFWERSRGIFLAPVFLVSPTLIGITTLATFLIOLERRKGVSSG 120

QY 137 IMTLFMVLVAVCALAIRSKIMTKLKDADVDLFRDITFYVYFSLIIQLVLCFSRSP 196
 DB 121 IMTLFMVLVAVCALAIRSKIMTKLKDADVDLFRDITFYVYFSLIIQLVLCFSRSP 180

QY 197 LFSETIHDPNCPPESSASPLSRITFMWITGLIVRGYQPLSGSLMLNKEDPSQVVPV 256
 DB 181 LFSETIHDPNCPPESSASPLSRITFMWITGLIVRGYQPLSGSLMLNKEDPSQVVPV 240

QY 257 LVNWKKECAKTRKQPKVYVSSKDPAPKSSKVDANEEVAILIVSPKEMNPISLFKV 316
 DB 241 LVNWKKECAKTRKQPKVYVSSKDPAPKSSKVDANEEVAILIVSPKEMNPISLFKV 300

QY 317 LYKTFGCVFLMSFFPKAIHDLMFSGQIILKLIKFNDRTPKADPMQGYFYTVLLFYVACL 376
 DB 301 LYKTFGCVFLMSFFPKAIHDLMFSGQIILKLIKFNDRTPKADPMQGYFYTVLLFYVACL 360

QY 377 QLTVLHGYEHCIFSGGRITKAVTGAVYRKALVTNSARKSSVGEIVNLSMDAQRMD 436
 DB 361 QLTVLHGYEHCIFSGGRITKAVTGAVYRKALVTNSARKSSVGEIVNLSMDAQRMD 420

QY 437 LATYINMWSAPLOVILALYLLMLNLGSPYLAGVAVWVLPVNAVWAMTKTKYQVAHMK 496
 DB 421 LATYINMWSAPLOVILALYLLMLNLGSPYLAGVAVWVLPVNAVWAMTKTKYQVAHMK 480

QY 497 SKDNRIKLMEIINGIKVLKLYAMELAFKDKVLAIRQEBELKVKKSAYLSAVGTFTWCT 556
 DB 481 SKDNRIKLMEIINGIKVLKLYAMELAFKDKVLAIRQEBELKVKKSAYLSAVGTFTWCT 540

QY 557 PELVALCTFAVYVTIDENNILLDAQTAVALPRLILFPLNLLPMVSSISVQASVSLKRL 616
 DB 541 PELVALCTFAVYVTIDENNILLDAQTAVALPRLILFPLNLLPMVSSISVQASVSLKRL 600

QY 617 RIFLSHELEPDSIERPVDGGTNSITVRNATFTWASRDPPTLNGITISIEGALVAV 676
 DB 601 RIFLSHELEPDSIERPVDGGTNSITVRNATFTWASRDPPTLNGITISIEGALVAV 660

QY 677 VGVGVGCKSSLLALLAEMDKVEGHVAIKGSVAVPOQAWI QNDSIRENILLFCQLEEPY 736
 DB 661 VGVGVGCKSSLLALLAEMDKVEGHVAIKGSVAVPOQAWI QNDSIRENILLFCQLEEPY 688

QY 737 YRSVIOACALLPDLLEIPSGDREIGKGVNLSSGQKQRYSLARAVYSNADIVLEDDPLS 796
 DB 689 YRSVIOACALLPDLLEIPSGDREIGKGVNLSSGQKQRYSLARAVYSNADIVLEDDPLS 721

QY 797 AYDAHVKGKHFENVIGPKMLKNKTRILLVTHSMSTYLPQVDVVIIVMSGKISMSGYOELL 856
 DB 797 AYDAHVKGKHFENVIGPKMLKNKTRILLVTHSMSTYLPQVDVVIIVMSGKISMSGYOELL 856

DB 722 AYDAHVKGKHFENVIGPKMLKNKTRILLVTHSMSTYLPQVDVVIIVMSGKISMSGYOELL 781
 QY 857 ARDGAFAEPIRTYASTQEDDAENRGYTGSGRKEKAKOMENGMULTDSAGKOLQROQLSS 916
 DB 782 ARDGAFAEPIRTYASTQEDDAENRGYTGSGRKEKAKOMENGMULTDSAGKOLQROQLSS 841

QY 917 SSSYSGDISHHNSTAELOKAEKKEETWMLMEADKQOTQOVLYSVYMDYMKALGLFTSF 976
 DB 842 SSSYSGDISHHNSTAELOKAEKKEETWMLMEADKQOTQOVLYSVYMDYMKALGLFTSF 901

QY 977 LSIFLFMCNVSALASNYWLSLWTDDEIVNGTOEHTKVRSLVYGALGISQGIAYFGSMA 1036
 DB 902 LSIFLFMCNVSALASNYWLSLWTDDEIVNGTOEHTKVRSLVYGALGISQGIAYFGSMA 961

QY 1037 VSIIGIILASRCLHVDLHSLIRBPMSEFETPSGNLVNRSKELDTYDSMIPETIKMPMG 1096
 DB 962 VSIIGIILASRCLHVDLHSLIRBPMSEFETPSGNLVNRSKELDTYDSMIPETIKMPMG 1021

QY 1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFVAVASSRDLKLESVSRSPVSHFN 1156
 DB 1022 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFVAVASSRDLKLESVSRSPVSHFN 1081

QY 1157 ETLIGVSVIRAFEOERFIHQSDLKVDENKAVYPSIVANRWLAVERECVNCIVLFAAL 1216
 DB 1082 ETLIGVSVIRAFEOERFIHQSDLKVDENKAVYPSIVANRWLAVERECVNCIVLFAAL 1141

QY 1217 FAVISRHSLSAGLVGSVSYSLQVTTYLMNLVMSSEMETNIIVAVERTKSESETERKAPW 1276
 DB 1142 FAVISRHSLSAGLVGSVSYSLQVTTYLMNLVMSSEMETNIIVAVERTKSESETERKAPW 1201

QY 1277 QIOETAPSSWPQVGRVEFPNVCYLRREDLDFLRHINVTINGEKVGIYGRGAGSSL 1336
 DB 1202 QIOETAPSSWPQVGRVEFPNVCYLRREDLDFLRHINVTINGEKVGIYGRGAGSSL 1261

QY 1337 TLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPDPVLFSSGLRNULPFSQYSD 1396
 DB 1262 TLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPDPVLFSSGLRNULPFSQYSD 1321

QY 1397 EBYWTSLELAHLKDPVSALPDKLDHECABEGENLSVGORQLVCLARALLRKTKILVDEA 1456
 DB 1322 EBYWTSLELAHLKDPVSALPDKLDHECABEGENLSVGORQLVCLARALLRKTKILVDEA 1381

QY 1457 TAAVDETDDLIOSTIRTOEDCTVLTIAHRLNTIMDYPRVYLDGELIOEGYAPSDLO 1516
 DB 1382 TAAVDETDDLIOSTIRTOEDCTVLTIAHRLNTIMDYPRVYLDGELIOEGYAPSDLO 1441

QY 1517 QRLGFYSMAKDAGLV 1531
 DB 1442 QRLGFYSMAKDAGLV 1456

RESULT 7
 O6UR05 CANPA
 ID O6UR05 CANPA PRELIMINARY; PRT; 1531 AA.
 AC O6UR05;
 DT 05-JUL-2004 (Tremblrel. 27, Created)
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE Multidrug resistance-associated protein 1.
 GN Name=MRP1;
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22404678; PubMed=12516967;
 RA Ma L., Pratt S.E., Cao J., Danczig A.H., Moore R.E., Slapak C.A.;
 RT "Identification and characterization of the canine multidrug
 resistance-associated protein.";
 RL Mol. Cancer Ther. 1:1335-1342(2002).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
RA Ma L., Prate S.E., Cao J., Danczig A.H., Moore R.E., Slapac C.A.;
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AY633728; AAC03148.1; -; mRNA.
DR Ensembl: ENSCAFG00000018208; Canis familiaris.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005524; F: ATP binding; IEA.
DR GO: GO:0016887; F: ATPase activity; IEA.
DR GO: GO:0042666; F: ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000166; F: nucleotide binding; IEA.
DR GO: GO:0006810; P: transport; IEA.
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR011527; ABC membrane 1.
DR InterPro: IPR001140; ABC TM transport.
DR InterPro: IPR003439; ABC transp_like.
DR InterPro: IPR005292; MRP_assoc.
DR InterPro: IPR000719; Prot_kinase.
DR Pfam: Pf00664; ABC_membrane; 2.
DR Pfam: Pf00005; ABC_tran; 2.
DR ProDom: PD000006; ABC transporter; 2.
DR SMART: SMO0382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS50929; ABC_TM1r; 2; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS0107; PROTEIN KINASE_ATP; UNKNOWN 1.
KM ATP-binding; Nucleotide-binding; Repeat; Transport.
SQ SEQUENCE 1531 AA; 171792 MW; 746361A71C6158BD CRC64;

Query Match 74.7%; Score 7272; DB 2; Length 1531;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;

QY 1 MALRGFSADSDSLMDMNTWNTNSPDFTCKONTUVMVPCRYLMACEFYEYLXSRH 60
DB 1 MALRGFSADSDSLPFWMDVSMNTSNPDFTCKONTUVMVPCRYLMACEFYEYLXSRH 60
QY 61 DRGIQMTPLNKRTALGFLIMVCMADLFYSFWEBSRGIFLAPVFLVSPFLIGITTLA 120
DB 61 DRGIQMTYINKRTALGFLVIMVCMADLFYSFWEBSRGKILAPVFLVSPFLIGITTLA 120
QY 121 TFLIQLEBRKGVSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLFEDITFYVYS 180
DB 121 TFLIQLEBRKGVSSGIMLTFWLVALCALILRSKIMTALKEDAIDVDFDVTFYVYS 180
QY 181 LLLIQVLVSCFSDSPFLFSETIHDNPPCESSASFLSITTFWMTGLIVRGYRQPLESD 240
DB 181 LVLQLVLVSCFSDSPFLFSETIHDNPPCESSASFLSVTFWMTGLIVRGYRQPLESTD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKKCACTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKKCAKSKROQRKITYSKDPKPKGSGQVDNBEAEVL 300
QY 301 IVKSPQKEWNSLFRKVLVYTFGPFYFLMSFFPKAIDHLMFSGPOLIKLILFVNDYAPD 360
DB 301 IVKTPQKERBSPLFRKVLVYTFGPFYFLMSFFKALHDLMMFAGPEILIKLILFVNDKAPD 360
QY 361 WQGFYFVTLFVTAQLQTLVHOYFHICFVSQMRKTAIVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYFVTLFVTAQLQTLVHOYFHI CFVSQMRKTAIVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYVNLNSVDQRFMDLATYINMIWSAPLOVILALYLLMLNLGSPVLGAVVMTLMPVN 480
DB 421 GEIYVNLNSVDQRFMDLATYINMIWSAPLOVILALYLLMLNLGSPVLGAVVMTLMPVN 480
QY 481 AVMAAMKTQYVVAHMKSKDNRIKLMNEILNGIKVLYLWAMELAFKQVLAIROBELKYLK 540
DB 481 AVMAAMKTQYVVAHMKSKDNRIKLMNEILNGIKVLYLWAMELAFKQVLAIROBELKYLK 540
QY 541 KSAIYLSAVGTTCVPLVALCTFAVYVYIDENNIIIDAQTAFFVSLAFNLIRPLANTLP 600
DB 541 KSAIYLSAVGTTCVPLVALCTFAVYVYIDENNIIIDAQTAFFVSLAFNLIRPLANTLP 600

QY 601 MVISSIVQASVSKRLNIFLSHELEBDSIERBPVXDGGSNTSITVNAATFTWARSDDPT 660
DB 601 MVISSIVQASVSKRLNIFLSHELEBDSIERBPVXDGGSNTSITVNAATFTWARSDDPT 660
QY 661 LNGTHFEPGALVAVVGVQCCGKSSLSALLAEMDKVEGVAKSGVAVVPOQAWIQND 720
DB 661 LNGTHFEPGSIYVAVVGVQCCGKSSLSALLAEMDKVEGVAKSGVAVVPOQAWIQND 720
QY 721 SLENNILFGQLPEEPYRSYIOACALLPDEIIPSGDRTEIGEGKVLSSGQKORVSLAR 780
DB 721 SLENNILFGQLPEEPYRSYIOACALLPDEIIPSGDRTEIGEGKVLSSGQKORVSLAR 780
QY 781 AVYSNADLYLFDPLSAVDHAKHILENVIIGPKMLKNKTRILVTHSMVLPQVDVIY 840
DB 781 AVYCDSDIYLFDPPLSAVDHAKHILENVIIGPKMLKNKTRILVTHSIYLPQVDVIY 840
QY 841 MSGKISMSQYQELARDAFAFLRTYASTEOQDAENGVTGVSGPKKEAKOMNGM 900
DB 841 MTGKISMSQYQELARDAFAFLRTYASTEOQDAENGVTGVSGPKKEAKOMNGM 900
QY 901 LVTDASGKOLOROLSSSSYSGDISRHNSYAELOKAEAKKEFTWKLMEADKAQTGVKL 960
DB 901 LVTDVAGKOLOROLSSSSYSGDISRHNSYAELOKAEAKKEFTWKLMEADKAQTGVKL 960
QY 961 SYVDYWKALGLFISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTKRLSVYG 1020
DB 961 SYVDYWKALGLFISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQGIANFGYSMAVSIIGIILASRCLHYDLHSILRSMSPFERIPSGNLVNRSEKL 1080
DB 1021 ALGISQGITFGYSMAVSIIGIFASRLHYDLHSILRSMSPFERIPSGNLVNRSEKL 1080
QY 1081 DTVDMSIPEVIKPMGSLFNVIAGCIYTLATPIAIIIPPLGIYFVQRFYVASSROL 1140
DB 1081 DTVDMSIPEVIKPMGSLFNVIAGCIYTLATPIAIIIPPLGIYFVQRFYVASSROL 1140
QY 1141 KRLESVRSVYSHFNETLLGVSVIRAFEOERFIHQSDUKVDENOKAYPSIVANRMLA 1200
DB 1141 KRLESVRSVYSHFNETLLGVSVIRAFEOERFIHQSDUKVDENOKAYPSIVANRMLA 1200
QY 1201 VRLCEVNCIVLPALFAVTSRHSLSAGLVGSYSLOYTTYINMLVRRSMSEETNIVA 1260
DB 1201 VRLCEVNCIVLPALFAVTSRHSLSAGLVGSYSLOYTTYINMLVRRSMSEETNIVA 1260
QY 1261 VERLKEYSETEKAPMIOETAPSSWPQYGRVFRYRCLYREDDLPVLRHNTVINGG 1320
DB 1261 VERLKEYSETEKAPMIOETAPSSWPQYGRVFRYRCLYREDDLPVLRHNTVINGG 1320
QY 1321 EKVGIVRTGAGKSLTLGLFRINESAEGEIIIDGINIAKITGHLDRFKITIIIPQDVLV 1380
DB 1321 EKVGIVRTGAGKSLTLGLFRINESAEGEIIIDGINIAKITGHLDRFKITIIIPQDVLV 1380
QY 1381 SGLSRNMLDPPSOYSDDEEVTSLSLAHLKDPVSALPKLDHECAEGENISVGOQOLVCL 1440
DB 1381 SGLSRNMLDPPSOYSDDEEVTSLSLAHLKDPVSALPKLDHECAEGENISVGOQOLVCL 1440
QY 1441 ARAILRRTKILVDEAFAVLDLTDLLIOSTIRQFDDCTVLTIAHRLNTIMDTRYIVL 1500
DB 1441 ARAILRRTKILVDEAFAVLDLTDLLIOSTIRQFDDCTVLTIAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIQEYGAPSDLQQRGLFYSMAXDAGLV 1531
DB 1501 DKGEIRECGQPSDLQQRGLFYSMAXDAGLV 1531
RESULT 8
059G19 HUMAN
ID 059G19 HUMAN PRELIMINARY; PRT; 1439 AA.
AC 059G19
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)

DE ATP-binding cassette, sub-family C, member 1 isoform 1 variant
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN
[1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain.
RA Tsuchi Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.,
RT "None Title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB209120; BAD92357.1; -; mRNA.
DR Ensemble; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR03593; AAA_ATPase.
DR InterPro; IPR01527; ABC membrane 1.
DR InterPro; IPR001140; ABC_TM_transport.
DR InterPro; IPR003439; ABC_TM_transport.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR00719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRfams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding.
FT
NON_TER 1
SEQUENCE 1439 AA; 160439 MW; 090D881ADBFB3077 CRC64;
Query Match 74.4%; Score 7242; DB 2; Length 1439;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1426; Conservative 0; Mismatches 1; Indels 10; Gaps 1;
QY 103 APVFLVSPPTLLIGITTLATFLIQERRKGVSSGIMLTFMVALVCAALILRSKIMTALK 162
DB 1 APVFLVSPPTLLIGITTLATFLIQERRKGVSSGIMLTFMVALVCAALILRSKIMTALK 60
QY 163 EDAGVDLFRDITTFVYFSLIIQVLVSCFSDRSPLFSETIHDNPPCESSASFLSRTTFW 222
DB 61 EDAGVDLFRDITTFVYFSLIIQVLVSCFSDRSPLFSETIHDNPPCESSASFLSRTTFW 120
QY 223 WITGLVIRGRQPLSGSDLSINKEDTSEQVPLVVKMKKECAKTRKOPKVVYSSKDP 282
DB 121 WITGLVIRGRQPLSGSDLSINKEDTSEQVPLVVKMKKECAKTRKOPKVVYSSKDP 180
QY 283 AQPKESSKVDANEEVEALIVKSPQKWNPSLFTKLYTFTGPFYFLMSFFFAIHDLMNFSG 342
DB 181 AQPKESSKVDANEEVEALIVKSPQKWNPSLFTKLYTFTGPFYFLMSFFFAIHDLMNFSG 240
QY 343 POLIKLILKFNVDNPKADNGCFYFTVLLFTTACIQTLVHQYFHTCVSGMRITAVIGA 402
DB 241 POLIKLILKFNVDNPKADNGCFYFTVLLFTTACIQTLVHQYFHTCVSGMRITAVIGA 300
QY 403 VYRKALVITNSARKSSTVGEIVNLMSYDAQRFMDLATYINMINSAPQVILALYMLNL 462
DB 301 VYRKALVITNSARKSSTVGEIVNLMSYDAQRFMDLATYINMINSAPQVILALYMLNL 360
QY 463 GPSVLAVAVMVLNVPYNAVMAKTKTYQVAHMSKDNRIKLNEIINGIKVLKYAMEL 522
DB 361 GPSVLAVAVMVLNVPYNAVMAKTKTYQVAHMSKDNRIKLNEIINGIKVLKYAMEL 420

QY 523 AFKDKVLAIRQBELKVLKKSAYLSAVGTFTWCTPFLVALCTFAVYVYTDENNILDAQTA 582
DB 421 AFKDKVLAIRQBELKVLKKSAYLSAVGTFTWCTPFLVALCTFAVYVYTDENNILDAQTA 480
QY 583 FVSLALFNILRFPNLILPMVYSSIVQASVSLKRLRIFLSHELEPDSIERRPVQGGGTN 642
DB 481 FVSLALFNILRFPNLILPMVYSSIVQASVSLKRLRIFLSHELEPDSIERRPVQGGGTN 540
QY 643 SITVRNATFTWASDPTLNGITFSIPFGALVAVGVGVGCKSSLSALLAEMDKVGHV 702
DB 541 SITVRNATFTWASDPTLNGITFSIPFGALVAVGVGVGCKSSLSALLAEMDKVGHV 600
QY 703 AIKGSVAVYVQQAIIQNDLSRENTLFGCQLLEPPYRSVITQACALLPLLEILPSGDRTEIG 762
DB 601 AIKGSVAVYVQQAIIQNDLSRENTLFGCQLLEPPYRSVITQACALLPLLEILPSGDRTEIG 660
QY 763 EKVNLSSGQKORVSLARAVYSNADIVLPDDPLSAVDAAHVKHIFENVIGPKMLKAKTR 822
DB 661 EKVNLSSGQKORVSLARAVYSNADIVLPDDPLSAVDAAHVKHIFENVIGPKMLKAKTR 720
QY 823 ILVTHSNSTYLPQVDVITVMSGGKISENGSYQELIARDGAFARFRTYASTBOEADAEN 881
DB 721 ILVTHSNSTYLPQVDVITVMSGGKISENGSYQELIARDGAFARFRTYASTBOEADAEN 780
QY 882 -----GVTVGSGGKRAKQENENMLVTDSAGKLOLQOLSSSSYSQSDISRHNSTA 932
DB 781 STWDEEBAVTVGSGGKRAKQENENMLVTDSAGKLOLQOLSSSSYSQSDISRHNSTA 840
QY 933 ELQKAKEKEETWKLMEADRAQTQVQLSYVMDYMKAIIGFISPLFLFMCHNVSALAS 992
DB 841 ELQKAKEKEETWKLMEADRAQTQVQLSYVMDYMKAIIGFISPLFLFMCHNVSALAS 900
QY 993 NYMISLMTDPIVNGQEHTRKRLSYVGAIGISQIAVFGYSMAVSIGGILASRCHVDL 1052
DB 901 NYMISLMTDPIVNGQEHTRKRLSYVGAIGISQIAVFGYSMAVSIGGILASRCHVDL 960
QY 1053 LHSILRSPMSFFERTPSGNLWRFSEKELDVDSNIPVIMFMSGLPNVIGACVILLAT 1112
DB 961 LHSILRSPMSFFERTPSGNLWRFSEKELDVDSNIPVIMFMSGLPNVIGACVILLAT 1020
QY 1113 PIAAIIIPPLGIYFFQRFYVASRQLKRLSESVSPVSHFNETLIGSVIRAFPEOE 1172
DB 1021 PIAAIIIPPLGIYFFQRFYVASRQLKRLSESVSPVSHFNETLIGSVIRAFPEOE 1080
QY 1173 RFIHQSDLKVDENQKAYPSIVANRWLAVERLCEVGNCTVLFALFAVISRHSLSAGLVGL 1232
DB 1081 RFIHQSDLKVDENQKAYPSIVANRWLAVERLCEVGNCTVLFALFAVISRHSLSAGLVGL 1140
QY 1233 SVSISLQVTTYLWMLVMSSEMETNIVAVRRLKEYSTEEKAPWQIOETAPSSWPQVGR 1292
DB 1141 SVSISLQVTTYLWMLVMSSEMETNIVAVRRLKEYSTEEKAPWQIOETAPSSWPQVGR 1200
QY 1293 VEFRRNYCLARYEDLDLFLRHINVTINGEKNVIGVGRTAGKSSLTGLFPIINSABEII 1352
DB 1201 VEFRRNYCLARYEDLDLFLRHINVTINGEKNVIGVGRTAGKSSLTGLFPIINSABEII 1260
QY 1353 IDGINIAKIGLHDLRFKRTIIPQDPVLFGSLRLMNLDPFSQYSDEEYVTSLELAHLKDFV 1412
DB 1261 IDGINIAKIGLHDLRFKRTIIPQDPVLFGSLRLMNLDPFSQYSDEEYVTSLELAHLKDFV 1320
QY 1413 SAPPDKLDHCCAGGENSELVSQRLQVCLARALARKKILVLDENATAVVDLETDLLIOSTI 1472
DB 1321 SAPPDKLDHCCAGGENSELVSQRLQVCLARALARKKILVLDENATAVVDLETDLLIOSTI 1380
QY 1473 RTQFEDCTVLTIAHRLNTIMDYTRVIVLQDKGEIOETGAPSDLLQOGLFYSMAKDGVLV 1531
DB 1381 RTQFEDCTVLTIAHRLNTIMDYTRVIVLQDKGEIOETGAPSDLLQOGLFYSMAKDGVLV 1439
RESULT 9
Q8HX05_BOVIN
ID Q8HX05_BOVIN PRELIMINARY; PRT; 1530 AA.
AC Q8HX05_

DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Multidrug resistance protein 1.
GN Name=MRP1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;
RA Taguchi Y., Saeiki K., Komano T.;
RT "Functional analysis of MRP1 cloned from bovine."
RL FEBS Lett. 521:211-213(2002).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB082124; BAC15550.1; -; mRNA.
DR HSSP; P08716; IMT0.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0062626; F:ATPase activity; coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane.1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRfams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1530 AA; 171666 MW; AABAF92ED7832703 CRC64;

Query Match 74.3%; Score 7230.5; DB 2; Length 1530;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;

QY 1 MALRGFCAGDSDDPLMDNMVMTNNSNPDTCQONTVLVWVPCFLMACEPPFLYLSRH 60
DB 1 MALNDPCSVDSDDLFWENNVMTNNSNPDTCQONTVLVWVPCFLMACEPPFLYLSHH 60

QY 61 DRGYIOMTPLNKTATLGFLLMIYVADLFYSFMRSRGIFLAIFYLVSPTLLGITTLLA 120
DB 61 DRGYIOMTHAKAKALGFLLMIYVADLFYSFMRSRGKLLAPFLVSPTLLGITTLLA 120

QY 121 TFLIQLERRKGVSSGIMLTFWLVALVCALATLRSKIMTALKEDAVDLFRDITFYVFS 180
DB 121 TFLIQLERRRGVSSGIMLTFWLALCALATLRSKIMTALKEDARVDVFRDITFYVFS 180

QY 181 LLLIQLVLSCTSDRPLFSETHDNPCESSASLSTITFWITGLIVRGROLESDD 240
DB 181 LVLQLVLSCTSDRPLFSETHDNPCESSASLSTITFWITGLVGMVQGRQLESSTD 240

QY 241 LMSLNKEDTSQVNVVLVKNMKKECAKTRKQPVKVVSCKDPAOPRESSKVDANEVEAL 300
DB 241 LMSLNKEDTSQVNVVLVKNMKKECAKTRKQPVKIVYSKDPARKSSKVDUNEVEARL 300

QY 301 IVKSPQKEMNSPLFVLYKTFGPFLMSFFFKAIHDLMMFSGPOLIKLIFKVDTKAPD 360
DB 301 IVKCPQKRDPDLFVLYKTFGPFLMSFLFKAVHDLMMFAGPEILKLLINVDNKAPE 360

QY 361 WQGYTYTLLPVTACLTQTLVLAHQYHICFVSGMRKTAVIGAVYKALVITMSAKSSIV 420
DB 361 WQGYTYTLLPVTACLTQTLVLAHQYHICFVSGMRKTAVIGAVYKALVITMAAKSSIV 420

QY 421 GEIVNLSVDAQRPMDIATYINMIWSAPLOVITALLYLMNLGSPVLAGVAVMLVMPVN 480
DB 421 GEIVNLSVDAQRPMDIATYINMIWSAPLOVITALLYLMNLGSPVLAGVAVMLVMPVN 480

QY 481 AVAMKTKTQVAMKSKDRIRIKMNEILNGIKYLKIYAMELAKDQVLAAROSELKYLK 540
DB 481 AVAMKTKTQVAMKSKDRIRIKMNEILNGIKYLKIYAMELAKDQVLAAROSELKYLK 540

QY 541 KSAIYASVGTFTWCTPFLVALCTFAVYVITDENNIIDAOAFASLAFNLRPEPLNLP 600
DB 541 KSAIYASVGTFTWCTPFLVALCTFAVYVITDENNIIDAOAFASLAFNLRPEPLNLP 600

QY 601 MVISSIVQASVSLKRLRIFLSHEELDPDSIORRPIKDAGATNSITVKNATFTMARNDPPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELDPDSIORRPIKDAGATNSITVKNATFTMARNDPPT 660

QY 661 LNGITFSIPGALVAVVGVCGGSSSILSALLAMDVBEVVALKGSVAAYVPOQAMOND 720
DB 661 LHGITFSVPEGSILVAVVGVCGGSSSILSALLAMDVBEVVALKGSVAAYVPOQAMONT 720

QY 721 SLRENIIFGCOLBEPYRSVIOACALLPDLLEILPSGRTIEGKGVNLGGQKORVSLAR 780
DB 721 SLRENIIFGRLQORRYKAVAEACALLPDLLEILPSGRTIEGKGVNLGGQKORVSLAR 780

QY 781 AVYSNADIYLPDDPLSAVDAAVHGKIFENVIGPKMLKNKTRILLVTHSMSTLPQVDVIV 840
DB 781 AVYCDSDVYLLDDPLSAVDAAVHGKIFENVIGPKMLKNKTRILLVTHSMSTLPQVDVIV 840

QY 841 MSGGKISEMGVQGLLARDGAFAEFLRTVASTBEQDAEENGAVGVSGPGKEAQMENG 900
DB 841 MSGGKISEMGVQGLLARDGAFAEFLRTVASTBEQDAEENGAVGVSGPGKEAQMENG 900

QY 841 MSGGKISEMGVQGLLARDGAFAEFLRTVASTBEQDAEENGAVGVSGPGKEAQMENG 900
DB 841 MSGGKISEMGVQGLLARDGAFAEFLRTVASTBEQDAEENGAVGVSGPGKEAQMENG 900

QY 901 LVTSAGKOLOROSSSSSYSGDISRRHNSYAELOKAKEETWKLMEADKATQGVKL 960
DB 901 LVTTAGKQKOROLSSSSSYSDVSOHHTSTAEARK-PGPEETWKLMEADKATQGVKL 959

QY 961 SVYDYMKAIGLFTSIFLFCMCHVASALASNYMSLMTDPIVNGTQHTKVLRSYVG 1020
DB 961 SVYDYMKAIGLFTSIFLFCMCHVASALASNYMSLMTDPIVNGTQHTKVLRSYVG 1019

QY 1021 ALGISQGIYVGYMAVSIIGIILASRCLHDLHLSILSPMSFPERPSSGULNRFSEL 1080
DB 1021 ALGISQGIYVGYMAVSIIGIILASRCLHDLHLSILSPMSFPERPSSGULNRFSEL 1079

QY 1081 DTVDSMTPEVIMKMGSLFNIVIGACIVILLATPIAIIIPPLGIIYFVQGFYVASSRQL 1140
DB 1081 DTVDSMTPEVIMKMGSLFNIVIGACIVILLATPIAIIIPPLGIIYFVQGFYVASSRQL 1139

QY 1141 KRLSEVSRSPPVSHNETLLGVSIVIRAEQGERFIHOSDLKVDENQKAYPSIVANRWLA 1200
DB 1141 KRLSEVSRSPPVSHNETLLGVSIVIRAEQGERFIHOSDLKVDENQKAYPSIVANRWLA 1199

QY 1201 VRLCEVNCIYLPALPAVVISRHSLSAGLVGLSVSYLSQVTTYLNMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALPAVVISRHSLSAGLVGLSVSYLSQVTTYLNMLVRMSSEMETNIVA 1259

QY 1261 VERLKEVSETEKAPWQIOETAPPSMPQVRVEFRNYCLYREDLDVLVHINVTING 1320
DB 1261 VERLKEVSETEKAPWQIOETAPPSMPQVRVEFRNYCLYREDLDVLVHINVTING 1319

QY 1321 EKVGVIGRTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIIPDPVLF 1380
DB 1321 EKVGVIGRTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIIPDPVLF 1379

QY 1381 SGLSRMNLDPFSQYSDSEEVWTSLELAHKDFVSALPDLKHECAEGGENLSVGQROLVCL 1440
DB 1381 SGLSRMNLDPFSQYSDSEEVWTSLELAHKDFVSALPDLKHECAEGGENLSVGQROLVCL 1439

QY 1441 ARAALRKTIIIVLBEATAVDLETDLLQSTIRIQFQEDCYTLTAHRLNTIMDYRYIVL 1500
DB 1441 ARAALRKTIIIVLBEATAVDLETDLLQSTIRIQFQEDCYTLTAHRLNTIMDYRYIVL 1499

QY 1501 DKGEIQEYGAPSDLLQGRGLFYMAKQAGLV 1531

Db 1500 DKGEIQEWGSPDLLQQRGLFYMAKDSGLV 1530

RESULT	10
Q9UQ98	HUMAN
ID	Q9UQ98_HUMAN
PRELIMINARY;	
PRT;	1400 AA.

CC -1- SIMILARITY: Belongs to the ABC transporter family.

DR EMBL, AF022824; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022825; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022829; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022829; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022827; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022830; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022832; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022834; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022845; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022844; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022843; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022842; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022841; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022838; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022837; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022836; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022835; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022852; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022851; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022850; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022849; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022848; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022847; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022846; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022833; AAB83982.1, JOINED, Genomic DNA.
DR EMBL, AF022831; AAB83982.1, JOINED, Genomic DNA.
DR HSSP, P08716; IMPT, Homo sapiens.
DR Ensemble, ENSG00000103222; Homo sapiens.
DR GO, GO:0016021; C.integral to membrane, NAS.
DR GO, GO:0004409; F1monocoatolate hydratase activity, NAS.
DR GO, GO:0006010; P.ribose synthase activity, NAS.
DR InterPro, IPR003593; AAA ATPase.
DR InterPro, IPR011527; ABC transpase.
DR InterPro, IPR001149; ABC TM transp.
DR InterPro, IPR003430; ABC transp.
DR InterPro, IPR005232; MRP_assoc.
DR InterPro, IPR007719; Prot_kinase.
DR Pfam, PF00664; ABC membrane, 2.
DR Pfam, PF00005; ABC tran, 1.
DR ProDom, PD000006; ABC transporter, 1.
DR SMART, SM00382; AAA; MRP.
DR TIGRfam, TIGR00957; MRP_assoc_prot, 1.
DR PROSITE, PS50929; ABC_TM1F, 2.
DR PROSITE, PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE, PS50893; ABC_TRANSPORTER_2; 2.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1
KW ATP-binding; Nucleotide-binding.

```

SQ      SEQUENCE 1400 AA; 157211 MW; C7C4DBB78DADDAAA CRC64;

```

Query Match	73.0%	Score 7104.5;	DB 2;	length 1400;
Best Local Similarity	92.4%;	Pred. NC. 0;		
Matches 1400;	Conservative 0;	Mismatches 0;	Indels 115;	Gaps 1.

QY	17	IMNVTNMTNSNDFPKFCQNTVLVWVPCYLMACPPFELIYLSRDRGVIQWTPINKTKTA	76
Db	1	DMNVTNMTNSNDFPKFCQNTVLVWVPCYLMACPPFELIYLSRDRGVIQWTPINKTKTA	60
QY	77	LGFLMTIWCNADLVYSEWERSRGIFLAPVPLVSEPTLLGITLTLATPFLIQLERRKGVOSG	136
Db	61	LGFLMTIWCNADLVYSEWERSRGIFLAPVPLVSEPTLLGITLTLATPFLIQLERRKGVOSG	120
QY	137	IMLTFWLVALVCAIIRSKIMTALKEDAOVDLFRDITFYVYFSLLLIQVLSCFSDRSP	196
Db	121	IMLTFWLVALVCAIIRSKIMTALKEDAOVDLFRDITFYVYFSLLLIQVLSCFSDRSP	180
QY	197	LFSEITHDPNCPRESSASFLSRTFMMITGILVNGVQPLGSGDLSMLNKEDTSEQVVPV	256
Db	181	LFSEITHDPNCPRESSASFLSRTFMMITGILVNGVQPLGSGDLSMLNKEDTSEQVVPV	240
QY	257	LYKMKKECKATRQOPKVVVSSKDPOPKESSVYDANEVEALIVKSPKEMNPSLFKV	316
Db	241	LYKMKKECKATRQOPKVVVSSKDPOPKESSVYDANEVEALIVKSPKEMNPSLFKV	300
QY	317	LYKTEGYPFLMSFFPKAIHDLMEGSGQILKLILKFVNDTKAPDMOGFYTVLLFVTA	376
Db	301	LYKTEGYPFLMSFFPKAIHDLMEGSGQILKLILKFVNDTKAPDMOGFYTVLLFVTA	360
QY	377	QTLVHLQYFHI CFVSGNRKITAVIGAYRRKALVITNSARKSSTYGEIYMLMSVDAORFMD	436
Db	361	QTLVHLQYFHI CFVSGNRKITAVIGAYRRKALVITNSARKSSTYGEIYMLMSVDAORFMD	420
QY	437	LATYINIMWSPLOVLIATLYLMLNLPSTVLAGAVWVLVAVNVAWMAKTKTYQVAHMK	496
Db	421	LATYINIMWSPLOVLIATLYLMLNLPSTVLAGAVWVLVAVNVAWMAKTKTYQVAHMK	480
QY	497	SKONRIKLMEILLNGIKVLKLYAMELAFKQVLAIRQEBLKVKKSAVLSAVGFTWVCT	556
Db	481	SKONRIKLMEILLNGIKVLKLYAMELAFKQVLAIRQEBLKVKKSAVLSAVGFTWVCT	540
QY	557	PFVLALCTFAYVYTTIDENNITLDACTAFVSLAFNILEPFLNILEPMVTISSIYQASVSLKRL	616
Db	541	PFVLALCTFAYVYTTIDENNITLDACTAFVSLAFNILEPFLNILEPMVTISSIYQASVSLKRL	600
QY	617	RIFLSHEELPDSIEREPVKDGGTNSITVRNMTFTMARBDPPLNITITESIPGALVAV	676
Db	601	RIFLSHEELPDSIEREPVKDGGTNSITVRNMTFTMARBDPPLNITITESIPGALVAV	660
QY	677	VGQVGCSSLSLSALLAEMDKVEGHVAIKGSVAVVPQAWI QNDSIRENLI,FGQLEBPY	736
Db	661	VGQVGCSSLSLSALLAEMDKVEGHVAI-----	688
QY	737	YRSVIOACALLPDLEILPSGDRTIEIGKGVNLGGOKQVRSLARAVYSNADIYLFDDPLS	796
Db	689	-----	688
QY	797	AVDAHVCKHIFENVIGKGMKNKTRTLLYTHMSYLPQVUVIIVMGGGKTSBMSYOBLL	856
Db	689	-----KTRLLVTHMSYLPQVUVIIVMGGGKTSBMSYOBLL	725
QY	857	ARDGAFAEFLRTYASTEQEOBDAEENGVTYVSGGPKAKQKQENGMVLVYDSAGKOLOROLSS	916
Db	726	ARDGAFAEFLRTYASTEQEOBDAEENGVTYVSGGPKAKQKQENGMVLVYDSAGKOLOROLSS	785
QY	917	SSSYSGDI SHHNSSTAELOKAEAKGEETWKLMEADKQTCQVUKLSVYWDYMKAI,GLFISF	976
Db	786	SSSYSGDISHHNSSTAELOKAEAKGEETWKLMEADKQTCQVUKLSVYWDYMKAI,GLFISF	845

```

QY 977 LSLFLMCMNHSALASNYWLSLMTDDPIVNGTOEHTKRLSVYGALGISOGIAVFGSMA 1036
Db 846 LSLFLMCMNHSALASNYWLSLMTDDPIVNGTOEHTKRLSVYGALGISOGIAVFGSMA 905
QY 1037 VSIIGIIASRGLHLDHLSIRSPMSFEFRTPSGNLVNRPEKEDTUDSMIPEYIKPMWG 1096
Db 906 VSIIGIIASRGLHLDHLSIRSPMSFEFRTPSGNLVNRPEKEDTUDSMIPEYIKPMWG 965
QY 1097 SLFNVIAGCIYLLATPIAIIIPPLGIYFVORFYVASSRQKRLSVSPVSHFN 1156
Db 966 SLFNVIAGCIYLLATPIAIIIPPLGIYFVORFYVASSRQKRLSVSPVSHFN 1025
QY 1157 ETLIGSVYTRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLAVERECVNCIYLPAL 1216
Db 1026 ETLIGSVYTRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLAVERECVNCIYLPAL 1085
QY 1217 FAVISRSISLNGVLSVSLQVTTYNNMLVRMSSEMETNIIVANERLKEVSEKEKAPW 1276
Db 1086 FAVISRSISLNGVLSVSLQVTTYNNMLVRMSSEMETNIIVANERLKEVSEKEKAPW 1145
QY 1277 QIOETAPPSWPQVGRVFRNYCLRYREDLDFVLHINVTNGEKGVIIGRTGAGKSSL 1336
Db 1146 QIOETAPPSWPQVGRVFRNYCLRYREDLDFVLHINVTNGEKGVIIGRTGAGKSSL 1205
QY 1337 TLGLFRINESAGEIIIDGINIAKIGHDLRPKITIIPODPVLSGSLRNNLDPPSQYSD 1396
Db 1206 TLGLFRINESAGEIIIDGINIAKIGHDLRPKITIIPODPVLSGSLRNNLDPPSQYSD 1265
QY 1397 EEWVTSLELAHLKOPVSLPKLDHECAEGENLSVGRQVCLARALLRTKILVDEA 1456
Db 1266 EEWVTSLELAHLKOPVSLPKLDHECAEGENLSVGRQVCLARALLRTKILVDEA 1325
QY 1457 TAAVDLETDLDIQTIRTOFEDCTVLTAAHRLNTIMDTRYVLVDKGEIOEGAPSDILO 1516
Db 1326 TAAVDLETDLDIQTIRTOFEDCTVLTAAHRLNTIMDTRYVLVDKGEIOEGAPSDILO 1385
QY 1517 QRGLEFYSMAKXAGLV 1531
Db 1386 QRGLEFYSMAKXAGLV 1400

```

```

RESULT 11
MRP1 MOUSE STANDARD; PRT; 1528 AA.
ID MRP1_MOUSE
AC 035379;
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-
  family C member 1).
GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mrp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Muscle;
RX MEDLINE=96251691; PubMed=8649356;
RA Deeley R.G.;
RT "Structure and expression of the messenger RNA encoding the murine
  multidrug resistance protein, an ATP-binding cassette transporter.";
RL Mol. Pharmacol. 49:962-971(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Miyai K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

```

```

RA Schriul L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbett L.E., Cousins S.,
RA Dalla B., Dzagani T.A., Fletcher C.F., Foxrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglot D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Naashima T., Numa K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hironaka-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Maki K., Kawai J., Aizawa K., Aikawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
CC -1- FUNCTION: May participate directly in the active transport of
  drugs into subcellular organelles or influence drug distribution
  indirectly (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
CC -1- SIMILARITY: Contains 2 ABC transporter domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use as long as its content is in no way modified and this statement is not
  removed.
CC -----
DR EMBL; AF022908; AAB80938.1; -; mRNA.
DR EMBL; AK029876; BAC26654.1; -; mRNA.
DR HSSP; P08716; IMT0.
DR Ensembl; ENSMUSG00000023088; Mus musculus.
DR MGI; MGI:102676; Abcc1.
DR GO; GO:0005887; C:integral to plasma membrane; IDA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR004339; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR PRODOM; PRO00006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMif; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Glycoprotein; Nucleotide-binding; Repeat; Transmembrane;
  Transport.
FT TOPO_DOM 1 33 Extracellular (By similarity).
FT TRANSMEM 34 54 1 (By similarity).
FT TOPO_DOM 55 74 Cytoplasmic (By similarity).
FT TRANSMEM 75 95 2 (By similarity).
FT TOPO_DOM 96 100 Extracellular (By similarity).
FT TRANSMEM 101 121 3 (By similarity).
FT TOPO_DOM 122 133 Cytoplasmic (By similarity).
FT TRANSMEM 134 154 4 (By similarity).
FT TOPO_DOM 155 172 Extracellular (By similarity).
FT TRANSMEM 173 193 5 (By similarity).
FT TOPO_DOM 194 317 Cytoplasmic (By similarity).
FT TRANSMEM 318 338 6 (By similarity).
FT TOPO_DOM 339 364 Extracellular (By similarity).

```

FT TRANSMEM 365 385 7 (By similarity).
FT TOPO_DOM 386 441 Cytoplasmic (By similarity).
FT TRANSMEM 442 462 8 (By similarity).
FT TOPO_DOM 463 465 Extracellular (By similarity).
FT TRANSMEM 466 486 9 (By similarity).
FT TOPO_DOM 487 548 Cytoplasmic (By similarity).
FT TRANSMEM 549 569 10 (By similarity).
FT TOPO_DOM 570 591 Extracellular (By similarity).
FT TRANSMEM 592 612 11 (By similarity).
FT TOPO_DOM 613 963 Cytoplasmic (By similarity).
FT TRANSMEM 964 984 12 (By similarity).
FT TOPO_DOM 985 1022 Extracellular (By similarity).
FT TRANSMEM 1023 1043 13 (By similarity).
FT TOPO_DOM 1044 1086 Cytoplasmic (By similarity).
FT TRANSMEM 1087 1107 14 (By similarity).
FT TOPO_DOM 1108 1108 Extracellular (By similarity).
FT TRANSMEM 1109 1129 15 (By similarity).
FT TOPO_DOM 1130 1200 Cytoplasmic (By similarity).
FT TRANSMEM 1201 1221 16 (By similarity).
FT TOPO_DOM 1222 1223 Extracellular (By similarity).
FT TRANSMEM 1224 1244 17 (By similarity).
FT TOPO_DOM 1245 1528 Cytoplasmic (By similarity).
FT DOMAIN 326 609 ABC transmembrane type-1.1.
FT DOMAIN 644 868 ABC transmembrane type-1.2.
FT DOMAIN 971 1253 ABC transporter 2.
FT DOMAIN 1290 1524 ATP 1 (Potential).
FT NE_BIND 678 685 ATP 2 (Potential).
FT NE_BIND 1324 1331 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 19 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1003 1003 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1528 AA; 171185 MW; 68FD1367D61DBB CRC64;

Query Match 71.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;

Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGCSADGSDPLMDMNTNTSNPDFTKCFONTLVWVPCGYLWACPFYFLYSRH 60
DB 1 MALRSCSADGSDPLMDMNTNTSNPDFTKCFONTLVWVPCGYLWACPFYFLYSRH 60
QY 61 DRGYIQMTPLNKTKTALGFLMTVCWADLFYSFEMERSGRIFLAPVPLVSEPLLGITLLA 120
DB 61 DRGYIQMTPLNKTKTALGFLMTVCWADLFYSFEMERSGRIFLAPVPLVSEPLLGITLLA 120
QY 121 TELIOERRKGVSSGMLTFPLVALCALAIRSKMTLAKEDAOVLDTRDTFTFYRES 180
DB 121 TELIOERRKGVSSGMLTFPLVALCALAIRSKMTLAKEDAOVLDTRDTFTFYRES 180
QY 121 TELIOERRKGVSSGMLTFPLVALCALAIRSKMTLAKEDAOVLDTRDTFTFYRES 180
DB 121 TELIOERRKGVSSGMLTFPLVALCALAIRSKMTLAKEDAOVLDTRDTFTFYRES 180
QY 181 LLLIQVLSCFSRSPLESETIHDNPNCPESSASFLSRIFFWMTGLIYVGYRQPLEGSD 240
DB 181 LLLIQVLSCFSRSPLESETIHDNPNCPESSASFLSRIFFWMTGLIYVGYRQPLEGSD 240
QY 181 LLLIQVLSCFSRSPLESETIHDNPNCPESSASFLSRIFFWMTGLIYVGYRQPLEGSD 240
DB 181 LLLIQVLSCFSRSPLESETIHDNPNCPESSASFLSRIFFWMTGLIYVGYRQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPVKVVYS-SKDPAPKSSSKVDANEVEA 299
DB 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPVKVVYS-SKDPAPKSSSKVDANEVEA 299
QY 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPVKVVYS-SKDPAPKSSSKVDANEVEA 299
DB 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPVKVVYS-SKDPAPKSSSKVDANEVEA 299
QY 300 LIYKSSQKEMNBSLFKYLKTPGPYFLMSFFFAIHDLMFSGPQIILKIKVNDTKAP 359
DB 300 LIYKSSQKEMNBSLFKYLKTPGPYFLMSFFFAIHDLMFSGPQIILKIKVNDTKAP 359
QY 301 LIYKSSQKEMNBSLFKYLKTPGPYFLMSFFFAIHDLMFSGPQIILKIKVNDTKAP 359
DB 301 LIYKSSQKEMNBSLFKYLKTPGPYFLMSFFFAIHDLMFSGPQIILKIKVNDTKAP 359
QY 360 DMGGFYFYLLEFYTACIOTLVHOFYHICVSGMRTKTAIVIGAVYKAVITTSARSSST 419
DB 360 DMGGFYFYLLEFYTACIOTLVHOFYHICVSGMRTKTAIVIGAVYKAVITTSARSSST 419
QY 361 DMGGFYFYLLEFYTACIOTLVHOFYHICVSGMRTKTAIVIGAVYKAVITTSARSSST 420
DB 361 DMGGFYFYLLEFYTACIOTLVHOFYHICVSGMRTKTAIVIGAVYKAVITTSARSSST 420
QY 420 VGSIVNLSVDAQRFDMLATYINMISAPLQVILALYLLMLNLGPSYLAGVAVMLVMPV 479
DB 420 VGSIVNLSVDAQRFDMLATYINMISAPLQVILALYLLMLNLGPSYLAGVAVMLVMPV 479
QY 421 VGSIVNLSVDAQRFDMLATYINMISAPLQVILALYLLMLNLGPSYLAGVAVMLVMPV 480
DB 421 VGSIVNLSVDAQRFDMLATYINMISAPLQVILALYLLMLNLGPSYLAGVAVMLVMPV 480
QY 480 NAYMAKTKTYQVAHKKSKDNRIKLMNEILNGIKTLVAMELAFQDKALAIQSEELKYL 539
DB 480 NAYMAKTKTYQVAHKKSKDNRIKLMNEILNGIKTLVAMELAFQDKALAIQSEELKYL 539
QY 481 NAYMAKTKTYQVAHKKSKDNRIKLMNEILNGIKTLVAMELAFQDKALAIQSEELKYL 540
DB 481 NAYMAKTKTYQVAHKKSKDNRIKLMNEILNGIKTLVAMELAFQDKALAIQSEELKYL 540
QY 540 KKSAYLSAVGTFTWVCTPFLVALCTFAVYVITDENNILLDAQTAFAVSLALFNILRFLNLI 599
DB 540 KKSAYLSAVGTFTWVCTPFLVALCTFAVYVITDENNILLDAQTAFAVSLALFNILRFLNLI 599

DB 541 KKSAYLSAVGTFTWVCTPFLVALCTFAVYVITDENNILLDAQTAFAVSLALFNILRFLNLI 600
QY 600 PMYISSIVQASVSLKRIKRIITLSHELEPDSIERRPVYKGGGNTSTYRNAFTWARSDDP 659
DB 601 PMYISSIVQASVSLKRIKRIITLSHELEPDSIERRPVYKGGGNTSTYRNAFTWARSDDP 659
QY 660 TLNGITFPIEGALVAVGVGGKSSLSALLAEMDKVGHVTLKSVAVVPOQAMIQN 719
DB 660 TLNGITFPIEGALVAVGVGGKSSLSALLAEMDKVGHVTLKSVAVVPOQAMIQN 719
QY 720 DSIARENILFCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSGGQKQVSLA 779
DB 720 DSIARENILFCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSGGQKQVSLA 779
QY 780 RAVYSNADYLPDDPLSAVDAAVGHKIFENVIGPKMKLKKTKRILVTHSMYSYLPQVDYII 839
DB 780 RAVYSNADYLPDDPLSAVDAAVGHKIFENVIGPKMKLKKTKRILVTHSMYSYLPQVDYII 839
QY 840 VMSGKISEMGSVOELLAPDGAFAELRTASTEQEDDAENGVTCVSGGKSKAKOMENG 899
DB 840 VMSGKISEMGSVOELLAPDGAFAELRTASTEQEDDAENGVTCVSGGKSKAKOMENG 899
QY 900 MLVTDGAKOLQROLSSSSYSGDISRRHNSSTAELOKAEKKEETWKLMEADRAQOTGOVY 959
DB 900 MLVTDGAKOLQROLSSSSYSGDISRRHNSSTAELOKAEKKEETWKLMEADRAQOTGOVY 959
QY 956 LSYVMDYKRAIGLFIPLFSLIFLPMCHNSALASNYMLSLMTDD-PIYNGIOEHTKRLSV 1018
DB 956 LSYVMDYKRAIGLFIPLFSLIFLPMCHNSALASNYMLSLMTDD-PIYNGIOEHTKRLSV 1018
QY 1019 YGALGISQGIYAVGYMAVSIIGIILASRCLADVLHSLILASPMSPFERITSGULVNRFSK 1078
DB 1019 YGALGISQGIYAVGYMAVSIIGIILASRCLADVLHSLILASPMSPFERITSGULVNRFSK 1078
QY 1079 ELDTVDSMTPEVIMKMGSLFENVYICIVILATPIAIIIPGLGIYFVQRPVYVASSR 1138
DB 1079 ELDTVDSMTPEVIMKMGSLFENVYICIVILATPIAIIIPGLGIYFVQRPVYVASSR 1138
QY 1136 QLRKLESVSRSPPYSHFNETLLGVSYIRAFEEQERFIHQSDLKVDENQKAYYSIVANRW 1195
DB 1136 QLRKLESVSRSPPYSHFNETLLGVSYIRAFEEQERFIHQSDLKVDENQKAYYSIVANRW 1195
QY 1199 LAVRLFECNGCYLPAALPAVIRSHSLAGLVLSYSYSIQVTTYLNLVYRSGSEMETNI 1258
DB 1199 LAVRLFECNGCYLPAALPAVIRSHSLAGLVLSYSYSIQVTTYLNLVYRSGSEMETNI 1258
QY 1259 VAVERLKEYSETKEAPMIOETAPPSMPOVGRVFRYKCLRYREDLDPVLHINVTIN 1318
DB 1259 VAVERLKEYSETKEAPMIOETAPPSMPOVGRVFRYKCLRYREDLDPVLHINVTIN 1318
QY 1319 GGEKVGIVRTGAGKSLTLGLFRINESAGBIIIDGINIAKIGLHDLRFKTIIPQDPV 1378
DB 1319 GGEKVGIVRTGAGKSLTLGLFRINESAGBIIIDGINIAKIGLHDLRFKTIIPQDPV 1378
QY 1379 LPSGSLRMLNDPSSQYSDSEEWTSLELAHLKDFVSLPDKLDHECAEGGENTSVGQROLV 1438
DB 1379 LPSGSLRMLNDPSSQYSDSEEWTSLELAHLKDFVSLPDKLDHECAEGGENTSVGQROLV 1438
QY 1439 CLABALLRKTKILVDEATPAVDLENDLLOSTIRQFQEDCTVLTAAHNLNTMDTRVY 1498
DB 1439 CLABALLRKTKILVDEATPAVDLENDLLOSTIRQFQEDCTVLTAAHNLNTMDTRVY 1498
QY 1499 VLDKGEIOEYGAPSDLQOORGLFYSWAKOAGLY 1531
DB 1499 VLDKGEIOEYGAPSDLQOORGLFYSWAKOAGLY 1531
QY 1496 VLDKGEVRECGAPSELLQORGIIFYSWAKOAGLY 1528
DB 1496 VLDKGEVRECGAPSELLQORGIIFYSWAKOAGLY 1528
RESULT 12
Q810E4 RAT PRELIMINARY; PRT; 1532 AA.
ID Q810E4 RAT
AC Q810E4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE ATP-binding cassette protein C1.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
ON NCBI_Taxid=10116;
[1]
NP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley, TISSUE=SpLeen;
RA Yabuuchi H., Takayanagi S., Ishikawa T.;
RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF487549; AAC05437.1; -; mRNA.
DR HSSP; P08716; IMT0.
DR Ensemble; ENSRNOG0000032748; Rattus norvegicus.
DR KGD; J112; Abcc1.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0006106; F:nucleotide binding; IEA.
DR GO; GO:000610; P:transport; IEA.
DR InterPro; IPR03593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldol/ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SMO0382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TM1F; 2.
DR PROSITE; PS50211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; UNKNOWN_1.
KM ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1532 AA; 171493 MW; 2E6393F63F5A3368 CRC64;

Query Match 71.2%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;

QY 1 MALAFGSGADSDPLMDMNTWNTSNPPTKCFONTLVVWPCFFYLAACFFPFFLYLSRH 60
DB 1 MALSSFCSSDSDPLMDMNTWNTSNPPTKCFONTLVVWPCFFYLAACFFPFFLYLSRH 60
QY 61 DRGIOMPLNKTATAGFLMTICWADLFYSFMRSRGIFLAAPFLVSPITLGIITLLA 120
DB 61 DRGIOMPLNKTATAGFLMTICWADLFYSFMRSRGIFLAAPFLVSPITLGIITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMVALVYCALALIRSKIMTALKEDAVDLEFDITFYVYS 180
DB 121 TFLIQLERRKGVSSGIMLTFMVALVYCALALIRSKIMTALKEDAVDLEFDITFYVYS 180
QY 121 TFLIQLERRKGVSSGIMLTFMVALVYCALALIRSKIMTALKEDAVDLEFDITFYVYS 180
DB 121 TFLIQLERRKGVSSGIMLTFMVALVYCALALIRSKIMTALKEDAVDLEFDITFYVYS 180
QY 181 LLLIQLVLSGDSPLSETIHDNPNCPRESSASLTSLITFMWITGLIVRGYRPLESSD 240
DB 181 LLLIQLVLSGDSPLSETIHDNPNCPRESSASLTSLITFMWITGLIVRGYRPLESSD 240
QY 181 LLLIQLVLSGDSPLSETIHDNPNCPRESSASLTSLITFMWITGLIVRGYRPLESSD 240
DB 181 LLLIQLVLSGDSPLSETIHDNPNCPRESSASLTSLITFMWITGLIVRGYRPLESSD 240
QY 241 LMSLNKEDTSBOVPLVIRKMKKECAKTRKQPVKVYVS-SKDPAPQPKSSKVDANEVEA 299
DB 241 LMSLNKEDTSBOVPLVIRKMKKECAKTRKQPVKVYVS-SKDPAPQPKSSKVDANEVEA 299
QY 241 LMSLNKEDTSBOVPLVIRKMKKECAKTRKQPVKVYVS-SKDPAPQPKSSKVDANEVEA 299
DB 241 LMSLNKEDTSBOVPLVIRKMKKECAKTRKQPVKVYVS-SKDPAPQPKSSKVDANEVEA 299
QY 300 LIVSPKSNPSPSLFKVLKTFGPFYLFMSFFKAIHIDLMFSGDILKLIKFVNDTRAP 359
DB 300 LIVSPKSNPSPSLFKVLKTFGPFYLFMSFFKAIHIDLMFSGDILKLIKFVNDTRAP 359
QY 301 LIVSGSHDRDPSLFKVLKTFGPFYLFMSFFKAIHIDLMFSGDILKLIKFVNDTRAP 360
DB 301 LIVSGSHDRDPSLFKVLKTFGPFYLFMSFFKAIHIDLMFSGDILKLIKFVNDTRAP 360
QY 360 DWGQYFTVLLFVTAACLTVLHQYFHCIVSGNRKIKTAIVGAVYRKALVITNSARKSST 419
DB 360 DWGQYFTVLLFVTAACLTVLHQYFHCIVSGNRKIKTAIVGAVYRKALVITNSARKSST 419

DB 361 DWGQYFTVLLFVTAACLTVLHQYFHCIVSGNRKIKTAIVGAVYRKALVITNSARKSST 420
QY 420 VGEIVNLSVDAQRFDMLATYINMWSAPLOVILATILMLNLPSTYAGAVANVLMV 479
DB 421 VGEIVNLSVDAQRFDMLATYINMWSAPLOVILATILMLNLPSTYAGAVANVLMV 480
QY 480 NAYVAMKTKTYQVAHMKSKNRIKLANEIINGIKVLTLYAMELAFKQKVALIROBELKVL 539
DB 481 NAYVAMKTKTYQVAHMKSKNRIKLANEIINGIKVLTLYAMELAFKQKVALIROBELKVL 540
QY 540 KKSATYLSAVGTFMVCPEPLVALCTFAVYVTTIBENNLIDQOTAFVSLATFNILRFPNITL 599
DB 541 KKSATYLSAVGTFMVCPEPLVALCTFAVYVTTIBENNLIDQOTAFVSLATFNILRFPNITL 600
QY 600 PMVSSIVQASVSLKRLIRFLSHELEPDSIERRPVQGGTNSITVRNATFTWASDDP 659
DB 601 PMVSSIVQASVSLKRLIRFLSHELEPDSIERRPVQGGTNSITVRNATFTWASDDP 660
QY 660 TLNGITSPGALVAVGVGCGKSLLSALLAEMDKVGHVAIKSVAVVPOQAMION 719
DB 661 TLNGITSPGALVAVGVGCGKSLLSALLAEMDKVGHVAIKSVAVVPOQAMION 720
QY 720 DSIARENTLPGQLEEPYRSVIOACALPDLEILPSGDRTEIGKGVNLSGGQKQORSIA 779
DB 721 DSIARENTLPGQLEEPYRSVIOACALPDLEILPSGDRTEIGKGVNLSGGQKQORSIA 780
QY 780 RAVYSNADIVLPDPLSAVDAAHVKHIFENVIGPKGMKKKTRILVTHSMSTYLPQVDVI 839
DB 781 RAVYSNADIVLPDPLSAVDAAHVKHIFENVIGPKGMKKKTRILVTHSMSTYLPQVDVI 840
QY 840 VMSGGKISEMGVQELLARDGAFEFRTYASTQEDDAENGVYSGGKKEAKQENB 899
DB 841 VMSGGKISEMGVQELLARDGAFEFRTYASTQEDDAENGVYSGGKKEAKQENB 900
QY 900 MLVTVDSAKOLORLSSSSYSGGDISRHNSTLELQAKBAKERTMTLMEADKQOTGOVK 959
DB 901 MLVTVDSAKOLORLSSSSYSGGDISRHNSTLELQAKBAKERTMTLMEADKQOTGOVK 959
QY 959 LSVYMDYMAKIGLIFSLIFLPMCNHVSALASNYWLSLWTD-PYVNGTOEHTKRLSV 1018
DB 960 LSVYMDYMAKIGLIFSLIFLPMCNHVSALASNYWLSLWTD-PYVNGTOEHTKRLSV 1019
QY 1019 YGALGISQIAVFGYSMAVSIIGILASRCLHVDLHSLNSPMSFFERTPSGNLVNRFSK 1078
DB 1020 YGALGISQIAVFGYSMAVSIIGILASRCLHVDLHSLNSPMSFFERTPSGNLVNRFSK 1079
QY 1079 ELDTVDMSIPVIMKMGSLFENVIGACIVILLAPRIAIIIPGLIYFPQRPYVASSR 1138
DB 1080 ELDTVDMSIPVIMKMGSLFENVIGACIVILLAPRIAIIIPGLIYFPQRPYVASSR 1139
QY 1139 QLRKLESVSRSPVYSHNETLLAGSVIRAFEEQERFIHQSDLKVDENOKAYVPSIVANRW 1198
DB 1140 QLRKLESVSRSPVYSHNETLLAGSVIRAFEEQERFIHQSDLKVDENOKAYVPSIVANRW 1199
QY 1199 LAVRLECGNCTVLPALPAVISHSISAGIVGLSVSYSLQVTTYLAWLVMSSEMETNI 1258
DB 1200 LAVRLECGNCTVLPALPAVISHSISAGIVGLSVSYSLQVTTYLAWLVMSSEMETNI 1259
QY 1259 VAVERLKEYSTKEAQMOTFAPPSSWPQVGVVEFRANCYARREDLDPLRHINVTIN 1318
DB 1260 VAVERLKEYSTKEAQMOTFAPPSSWPQVGVVEFRANCYARREDLDPLRHINVTIN 1319
QY 1319 GGEKVGIVGRTGACKSLTGLFPIINSAGEEIIIDGINIAKIGLHNLRFKITTIPDPV 1378
DB 1320 GGEKVGIVGRTGACKSLTGLFPIINSAGEEIIIDGINIAKIGLHNLRFKITTIPDPV 1379
QY 1379 LFSGSLRNMLDPFSQYSDDEEVTSLLAHLKDPVSALPDKIDHECAEGENLSVGQROLV 1438
DB 1380 LFSGSLRNMLDPFSQYSDDEEVTSLLAHLKDPVSALPDKIDHECAEGENLSVGQROLV 1439
QY 1439 CLARALLRKTKILVLDATAVAVDLETDLLQSTIRTOPEDCVTVTIARLNTINDYTRVI 1498
DB 1440 CLARALLRKTKILVLDATAVAVDLETDLLQSTIRTOPEDCVTVTIARLNTINDYTRVI 1499

Qy 1499 VLDKGEIOEYGAPSDILQORGLFYSMADAGLV 1531
Db 1500 VLDKGEIRECAPSBELOQRGVFTSMADAGLV 1532

RESULT 13

Q8CG09 RAT PRELIMINARY; PRT; 1532 AA.

AC Q8CG09_1 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Multidrug resistance-associated protein 1.
GN Name=Abcc1; Synonyms=Mrp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
NCBI_Taxid=10116;
RN [1]

NP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Yang Z., Li C.S.W., Shen D.D., Ho R.J.Y.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AY170916; AAA6532.1; -, mRNA.
DR HSSP: P08716; 1MT0.

DR Ensemble1; ENSRNOG0000032748; Rattus norvegicus.

DR RGD; 3112; Abcc1.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0016887; F:ATPase activity; IEA.

DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR011527; ABC_membrane_1.

DR InterPro; IPR001140; ABC_TM_transp.

DR InterPro; IPR003439; ABC_transp_like.

DR InterPro; IPR001395; Aldo/ket_red.

DR InterPro; IPR005292; MRP_assoc.

DR Pfam; PF00664; ABC_membrane; 2.

DR Pfam; PF00005; ABC_tran; 2.

DR Prodom; PD000006; ABC_transporter; 2.

DR SMART; SM00382; AAA; 2.

DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.

DR PROSITE; PS00929; ABC_TMIF; 2.

DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.

DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.

DR PROSITE; PS00063; ALDO-KETO REDUCTASE; 3; UNKNOWN_1.

DR ATP-binding; Nucleotide-binding; Repeat.

DR KW

DR SEQUENCE 1532 AA; 171558 MW; C18F3554D85732A CRC64;

Query Match 71.2%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1336; Conservative 100; Mismatches 92; Indels 3; Gaps 3;

Qy 1 MALRGFCADGSDPLMDMNTVNTSNPDFTKCFQNTLVVWPCFYLMACEPFYFLYSRH 60
Db 1 MALRSFCSGSDSDPLMDMNTVNTSNPDFTKCFQNTLVVWPCFYLMACEPFYFLYSRH 60
Qy 61 DRGTYQMTPLNKTITAGFLIMTYCADLFTSFMEGRSGIFLAFVLVSPLLGITTLLA 120
Db 61 DRGTYQMTNKAATAGFLIMTYCADLFTSFMEGRSGIFLAFVLVSPLLGITTLLA 120
Qy 121 TFLIQLEERRKGVSSGIMLTFMLVALYCALAIIISKIMTALKEDAVDLDFTDITFYVYFS 180
Db 121 TFLIQLEERRKGVSSGIMLTFMLVALYCALAIIISKIISALKKDAQVDMFRDASFFLYFT 180
Qy 181 LLLLIQVLVSCFSDRSLPSEFTIHDPNCPRESSASFLSRITFWWITGLIVGYRQPLGSD 240
Db 181 LVFIQVLVSCFSDRSLPSEFTVRDNPCESSASFLSRITFWWITGMVVOGYRQPLGSSD 240

Qy 241 LMSLNKEDTSROVYVLYVKNWKECAKTRKQPVKNVYS-SKDPAPQKSSKVDANEVEYA 299
Db 241 LMSLNKEDTSEEVVPLVNNMKKECVKSRKQPVPIVYAPDPKPKSSOLDVNEVEYA 300
Qy 300 LTVKSPKEMWNPSEFLKYLKTFGPFVLMSEFFFAKIHLMFSPGQIILKLFVNDTKAP 359
Db 301 LTVKSHKNDPSLFLKYLKTFGPFVLMSEFLYKALHLMFAGBEILELINFVNDRAP 360
Qy 360 DMQGYFTYVLLFVTACIQTLVHQYFHI CFVSGKRITKTAIVAGAYRRALVITNSARSKST 419
Db 361 DMQGYLTALLFVSACIQTLALHQYFHI CFVSGKRITKTAIVAGAYRRALVITNSARSKST 420
Qy 420 VGEIYVNMVSADQGFMDLATYINMNSAPLOVLTALVLMNINGPVSLAGAVNWLMPV 479
Db 421 VGEIYVNMVSADQGFMDLATYINMNSAPLOVLTALVLMNINGPVSLAGAVNWLMPV 480
Qy 480 NAWAMAKTKYQVAMHMSKDNRIKLAMEIINGIVLKLAYAMELAFKDKVLAIROBELKVL 539
Db 481 NAWAMAKTKYQVAMHMSKDNRIKLAMEIINGIVLKLAYAMELAFKDKVLAIROBELKVL 540
Qy 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITDENNTLDAQTAFLVSLAFNILRFPPLNT 599
Db 541 KKSAYLAAGVGTFTWCTPFLVALCTFAVYVITDENNTLDAKKAFLVSLAFNILRFPPLNT 600
Qy 600 PMVTSITVOASVSLKRLRIFLSHELEPPDSIERRPVVDGGGTNSITVRNATFTAKSDPP 659
Db 601 PMVTSITVOASVSLKRLRIFLSHELEPPDSIERRSIXDGGGMSITVKNATFTAKSDPP 660
Qy 660 TLNGITSIPEGALVAVVGVGCGKSLLSALAEPMKVEGHVAKGSVAVPOAMON 719
Db 661 TLNGITSIPEGALVAVVGVGCGKSLLSALAEPMKVEGHVAKGSVAVPOAMON 720
Qy 720 DSIARENILFCQLEPEYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGQKQVSLA 779
Db 721 DSIARENILFCQLEPEYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGQKQVSLA 780
Qy 780 RAYVSNADITLFPDDPSLVDAHVGHKIFENVITGPKMKLKNKTRILVTHSMSTLPQVDVIT 839
Db 781 RAYVSNADITLFPDDPSLVDAHVGHKIFENVITGPKMKLKNKTRILVTHSMSTLPQVDVIT 840
Qy 840 VMSGGKISEMGVYOELIARDAQFAEFLRTYASTBOEDAEENGVTGVSQPKKAKOMENG 899
Db 841 VMSGGKISEMGVYOELIARDAQFAEFLRTYASTBOEDAEENGVTGVSQPKKAKOMENG 900
Qy 900 MLVYDSAGKOLQROLSSSSSSGDISRHNSSTAELQKAEAKKETWLMADKAQTOGVK 959
Db 901 MLVYDSAGKOLQROLSSSSSSGDISRHNSSTAELQKAEAKKETWLMADKAQTOGVK 959
Qy 960 LSYVMDYMKAIQGLFISFLSIFLMCNHVSALASNYWLSLMTDD-PIYNGTQEHTRVLSV 1018
Db 960 LSYVMDYMKAIQGLFISFLSIFLMCNHVSALASNYWLSLMTDD-PIYNGTQEHTRVLSV 1019
Qy 1019 YGALGTSOGIYAVFGYSAVSIIGIILASRCHVDLHLSLSPMSFEERTSGNLYNFSK 1078
Db 1020 YGALGTSOGIYAVFGYSAVSIIGIILASRCHVDLHLSLSPMSFEERTSGNLYNFSK 1079
Qy 1079 ELDTVDSMIPEVAKMFNGSLFNVIYACTVILLATPIAIIIPRLGILYFFVQRPVYVASSR 1138
Db 1080 ELDTVDSMIPEVAKMFNGSLFNVIYACTVILLATPIAIIIPRLGILYFFVQRPVYVASSR 1139
Qy 1139 QLRLESVSRSRPSVYSHNETLLGVSVIRAEBOERFIHQSDLYVDENOKAYYSISVANRW 1198
Db 1140 QLRLESVSRSRPSVYSHNETLLGVSVIRAEBOERFIHQSDLYVDENOKAYYSISVANRW 1199
Qy 1199 LAVRLCEVGCITLFPALFAVISRHSLSAGLVGSYSIQVTTYLNWLVRMSSEMTNI 1258
Db 1200 LAVRLCEVGCITLFPALFAVISRHSLSAGLVGSYSIQVTTYLNWLVRMSSEMTNI 1259
Qy 1259 VAVERLKEYSETEKEAPWQIOETAPSSWQOVGRVVEFRNYCLARYRBDLVLRHINVTIN 1318
Db 1260 VAVERLKEYSETEKEASWQIOETAPSSWQOVGRVVEFRNYCLARYRBDLVLRHINVTIS 1319

QY 1319 GGEKVIIVGTGAGKSSITLGLFRINESAGEEIIIDGINIAKIGLHDIREFKIIITIPDDPV 1378
DB 1320 GGEKVIIVGTGAGKSSITLGLFRINESAGEEIIIDGINIAKIGLHDIREFKIIITIPDDPV 1379
QY 1379 LFGSLTLMNIDPFQSYDEEVTSLTSLAHKDPVSALPDKLDHCBAGGEMLSVGORGLV 1438
DB 1380 LFGSLTLMNIDPFQSYDEEVTSLTSLAHKDPVSALPDKLDHCBAGGEMLSVGORGLV 1439
QY 1439 CLARALKRTKIILVDEATAVDETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYREVI 1498
DB 1440 CLARALKRTKIILVDEATAVDETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYREVI 1499
QY 1499 VLDKGEIOEGYCBSDLLQORGLFYSMAKAGLV 1531
DB 1500 VLDKGEIRECGAPSELLQORGVFYSMAKAGLV 1532
RESULT 14
QY 0810G9_RAT PRELIMINARY; PRT; 1523 AA.
ID 0810G9_RAT
AC 0810G9;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE ATP-binding cassette protein C1 variant A.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Takayanagi S., Ichikawa T.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY174892; AAC44983.1; -, mRNA.
DR HSSP; P08716; 1MTO.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR01140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TM1F; 2; 1.
DR PROSITE; PS50211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1523 AA; 170505 MW; E40337051A1CB9C6 CRC64;
Query Match 70.8%; Score 6892; DB 2; Length 1523;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;
QY 1 MALGFCGADGSDPLMDNMVMTNSPPTKCFONTLVWVPCFYLMACFPYFLYLSRH 60
DB 1 MALGFCGADGSDPLMDNMVMTNSPPTKCFONTLVWVPCFYLMACFPYFLYLSRH 60
QY 61 DRGYIOMTPLNKXTALGFLMIWCMADLFYSFWERSRGIFLAPVFLVSPULLGITTLA 120

DB 61 DRGYIOMTHLNKATYALGFLMIWCMADLFYSFWERSRGIFLAPVFLVSPULLGITTLA 120
QY 121 TFLIQERRKGVSSGIMLTFMWALVACALAIIRSKIMTLAKEDAVDLDFEDITFYVYS 180
DB 121 TFLIQERRKGVSSGIMLTFMWALVACALAIIRSKIMTLAKEDAVDLDFEDITFYVYS 180
QY 181 LLLIQVLVSCFSDRSLPSETTHDNPCCPESSAFLSRTIFMWITGLIVRGYRPLEGSD 240
DB 181 LVFIQLVLSGFSDDSPLESETVRDNPCCPESSAFLSRTIFMWITGMVVGQRPLKSSD 240
QY 241 LMSLNKEDTSEVQVPLVYKMKKCATRKQPVNVYS -SDPAQPPKSSAVDANEVEA 299
DB 241 LMSLNKEDTSEVQVPLVYKMKKCATRKQPVNVYS -SDPAQPPKSSAVDANEVEA 299
QY 300 LMSLNKEDTSEVQVPLVYKMKKCATRKQPVNVYS -SDPAQPPKSSAVDANEVEA 300
DB 300 LMSLNKEDTSEVQVPLVYKMKKCATRKQPVNVYS -SDPAQPPKSSAVDANEVEA 300
QY 301 LMSLNKEDTSEVQVPLVYKMKKCATRKQPVNVYS -SDPAQPPKSSAVDANEVEA 300
DB 301 LMSLNKEDTSEVQVPLVYKMKKCATRKQPVNVYS -SDPAQPPKSSAVDANEVEA 300
QY 360 DMQGYFTVLLFVTACIQTLVLAHQYFHCYVSGNRKRTAVIYGAAYRKALVITNSARKST 419
DB 360 DMQGYFTVLLFVTACIQTLVLAHQYFHCYVSGNRKRTAVIYGAAYRKALVITNSARKST 419
QY 361 DMQGYFTVLLFVTACIQTLVLAHQYFHCYVSGNRKRTAVIYGAAYRKALVITNSARKST 420
DB 361 DMQGYFTVLLFVTACIQTLVLAHQYFHCYVSGNRKRTAVIYGAAYRKALVITNSARKST 420
QY 420 VGEIIVNLSVDAQRFMDLATYINMINSAPLOVILATFLMLNLGSPVLGAVNWLWVPV 479
DB 420 VGEIIVNLSVDAQRFMDLATYINMINSAPLOVILATFLMLNLGSPVLGAVNWLWVPV 479
QY 421 VGEIIVNLSVDAQRFMDLATYINMINSAPLOVILATFLMLNLGSPVLGAVNWLWVPV 480
DB 421 VGEIIVNLSVDAQRFMDLATYINMINSAPLOVILATFLMLNLGSPVLGAVNWLWVPV 480
QY 480 NAWNAMTKTYQVAMHMSKDNRIKLMNEILNGIKVLYLWAMELAFKQVLAIRQBELKVL 539
DB 480 NAWNAMTKTYQVAMHMSKDNRIKLMNEILNGIKVLYLWAMELAFKQVLAIRQBELKVL 539
QY 481 NAWNAMTKTYQVAMHMSKDNRIKLMNEILNGIKVLYLWAMELAFKQVLAIRQBELKVL 540
DB 481 NAWNAMTKTYQVAMHMSKDNRIKLMNEILNGIKVLYLWAMELAFKQVLAIRQBELKVL 540
QY 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYTTIDENMLDAQTAFVSLAFNLRRPMLNL 599
DB 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYTTIDENMLDAQTAFVSLAFNLRRPMLNL 599
QY 541 KKSAYLSAVGTFTWCTPFLVALCTFAVYTTIDENMLDAQTAFVSLAFNLRRPMLNL 600
DB 541 KKSAYLSAVGTFTWCTPFLVALCTFAVYTTIDENMLDAQTAFVSLAFNLRRPMLNL 600
QY 600 PMVTSIYQASVSLKRLRIFLSHEELPDSIERRPVYDGGGTSNITRANAFITARSPP 659
DB 600 PMVTSIYQASVSLKRLRIFLSHEELPDSIERRPVYDGGGTSNITRANAFITARSPP 659
QY 601 PMVTSIYQASVSLKRLRIFLSHEELPDSIERRPVYDGGGTSNITRANAFITARSPP 660
DB 601 PMVTSIYQASVSLKRLRIFLSHEELPDSIERRPVYDGGGTSNITRANAFITARSPP 660
QY 660 TLNGITFSPGALVAVVGVGCGKSSLSALLMEMDKVEGVAMIKGSVAVVPOQAMION 719
DB 660 TLNGITFSPGALVAVVGVGCGKSSLSALLMEMDKVEGVAMIKGSVAVVPOQAMION 719
QY 661 TLNGITFSPGALVAVVGVGCGKSSLSALLMEMDKVEGVAMIKGSVAVVPOQAMION 720
DB 661 TLNGITFSPGALVAVVGVGCGKSSLSALLMEMDKVEGVAMIKGSVAVVPOQAMION 720
QY 720 DLSRENILFGQLEPEYRYSYIQAALLPDLEILPSGDREIGEGVNLGGQKORVSLA 779
DB 720 DLSRENILFGQLEPEYRYSYIQAALLPDLEILPSGDREIGEGVNLGGQKORVSLA 779
QY 721 DLSRENILFGQLEPEYRYSYIQAALLPDLEILPSGDREIGEGVNLGGQKORVSLA 780
DB 721 DLSRENILFGQLEPEYRYSYIQAALLPDLEILPSGDREIGEGVNLGGQKORVSLA 780
QY 780 RAVYSNADYILFDDPLSANDAHVGHKIFENVYIGKMLKXKTRILVYTHSMGLYQVDYII 839
DB 780 RAVYSNADYILFDDPLSANDAHVGHKIFENVYIGKMLKXKTRILVYTHSMGLYQVDYII 839
QY 781 RAVYSNADYILFDDPLSANDAHVGHKIFENVYIGKMLKXKTRILVYTHSMGLYQVDYII 840
DB 781 RAVYSNADYILFDDPLSANDAHVGHKIFENVYIGKMLKXKTRILVYTHSMGLYQVDYII 840
QY 840 VMSGKTISEMGSYOELLARDGAPAEPLRTYASTBOEDAEENGVYGVSGPKKAKOMENG 899
DB 840 VMSGKTISEMGSYOELLARDGAPAEPLRTYASTBOEDAEENGVYGVSGPKKAKOMENG 899
QY 841 VMSGKTISEMGSYOELLARDGAPAEPLRTYASTBOEDAEENGVYGVSGPKKAKOMENG 900
DB 841 VMSGKTISEMGSYOELLARDGAPAEPLRTYASTBOEDAEENGVYGVSGPKKAKOMENG 900
QY 900 MLVYDSAGKQORQLSSSSSYSGDISRHNSSTAEIQRKAEAKKBTWKLMEADKAQTGVYK 959
DB 900 MLVYDSAGKQORQLSSSSSYSGDISRHNSSTAEIQRKAEAKKBTWKLMEADKAQTGVYK 959
QY 901 ILVYDVAQKPL-----HSVYTNQHSSTAEIQS -GVKEEFTWKLMEADKAQTGVYK 950
DB 901 ILVYDVAQKPL-----HSVYTNQHSSTAEIQS -GVKEEFTWKLMEADKAQTGVYK 950
QY 960 LSVYDVAQKPL-----HSVYTNQHSSTAEIQS -GVKEEFTWKLMEADKAQTGVYK 1018
DB 960 LSVYDVAQKPL-----HSVYTNQHSSTAEIQS -GVKEEFTWKLMEADKAQTGVYK 1018
QY 951 LSVYDVAQKPL-----HSVYTNQHSSTAEIQS -GVKEEFTWKLMEADKAQTGVYK 1010
DB 951 LSVYDVAQKPL-----HSVYTNQHSSTAEIQS -GVKEEFTWKLMEADKAQTGVYK 1010
QY 1019 YGALGISGCIAPVGYMAVSTIGTILASCLHVDLHSLSTRSPMSFEETPGSNLVNRPXK 1078
DB 1019 YGALGISGCIAPVGYMAVSTIGTILASCLHVDLHSLSTRSPMSFEETPGSNLVNRPXK 1078
QY 1011 YGALGISGCIAPVGYMAVSTIGTILASCLHVDLHSLSTRSPMSFEETPGSNLVNRPXK 1070
DB 1011 YGALGISGCIAPVGYMAVSTIGTILASCLHVDLHSLSTRSPMSFEETPGSNLVNRPXK 1070
QY 1079 ELDTVDWMIPEVIKFMGSLFNVGACIVILATPAAIIIPPIGLIYFFVQRFVYASSR 1138
DB 1079 ELDTVDWMIPEVIKFMGSLFNVGACIVILATPAAIIIPPIGLIYFFVQRFVYASSR 1138
QY 1071 ELDTVDWMIPEVIKFMGSLFNVGACIVILATPAAIIIPPIGLIYFFVQRFVYASSR 1130
DB 1071 ELDTVDWMIPEVIKFMGSLFNVGACIVILATPAAIIIPPIGLIYFFVQRFVYASSR 1130
QY 1139 QKRLSEVSRSVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRW 1198
DB 1139 QKRLSEVSRSVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRW 1198

Db 1131 QLKRLSESVSRSPVYSHENETLLAGVIRAFEEGRFRTQSDLDKVDENOKAYPSIVANRW 1190
QY 1199 LAARLECVGNCTIVLFAALFANISRHSLSAGLVGSVSQVOTTYTAMLVPMSEMETNI 1258
Db 1191 LAARLECVGNCTIVLFAALFANISRHSLSAGLVGSVSQVOTTYTAMLVPMSEMETNI 1250
QY 1259 VAVERLKEVESETEGEAPMOIOETAPSPSPGVGRVFRNVCYLRREDDFLRIINNTIN 1318
Db 1251 VAVERLKEVESETEGEASWQIOETAPSPSPGVGRVFRNVCYLRREDDFLRIINNTIN 1310
QY 1319 GGEKVGIVGRTRGACKSSLTGLFPRINSAEGEIIIDGINIAKIGLHDLRFKITTIIPODPV 1378
Db 1311 GGEKVGIVGRTRGACKSSLTGLFPRINSAEGEIIIDGINIAKIGLHDLRFKITTIIPODPV 1370
QY 1379 LFGSGLSMNLDPPSQYDEDEWVTSLELAHLKDPYSALPKLDHECAEGENLSVGORQLV 1438
Db 1371 LFGSGLSMNLDPPSQYDEDEWVTSLELAHLKDPYSALPKLDHECAEGENLSVGORQLV 1430
QY 1439 CLARALKRKTILVLDATAVADLETDDLOSTIRTOPEDCVTYLTIAHRLNTIMDYRVI 1498
Db 1431 CLARALKRKTILVLDATAVADLETDDLOSTIRTOPEDCVTYLTIAHRLNTIMDYRVI 1490
QY 1499 VLDKGEIOEYGAPEDLLQORGLFYSMKADGLV 1531
Db 1491 VLDKGEIOEYGAPEDLLQORGLFYSMKADGLV 1523

RESULT 15
05F364 CHICK PRELIMINARY; PRT: 1525 AA.
ID 05F364 CHICK PRELIMINARY; PRT: 1525 AA.
AC 05F364
DT 10-MAY-2005 (Tremblere1.30, Created)
DT 10-MAY-2005 (Tremblere1.30, Last sequence update)
DT 10-MAY-2005 (Tremblere1.30, Last annotation update)
DE Hypochemical protein.
GN ORFNames=RCJMB04.32d20;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Fiedler P., Kuter S., Blagoderaki A., Kostovska D., Kotler M.,
RA Plachy J., Carnuci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursa lymphocytes to facilitate
RT genefunction analysis";
RL Genome Biol. 6:R6-R6(2005).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB51786; CAH65420.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:nucleotide binding; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_transporter.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR Prodom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

KM ARP-binding; Hypochemical protein; Nucleotide-binding; Repeat;
KM Transport. 1525 AA; 170972 MW; 387246711B878FD CRC64;
SQ SEQUENCE 1525 AA; 170972 MW; 387246711B878FD CRC64;
Query Match 64.7%; Score 6301; DB 2; Length 1525;
Best Local Similarity 77.6%; Pred. No. 0;
Matches 1199; Conservative 176; Mismatches 148; Indels 20; Gaps 7;
QY 1 MALRFGSAGSDPDLMDNNTNNTSNDFPKCFRONTYLVVPCFFYLACGPFYLYSRH 60
1 MGIESLCSADASEPFMDNLTWHTENPDFQCQNTYLVVPCFFYLACGPFYLYSRH 60
Db 1 MGIESLCSADASEPFMDNLTWHTENPDFQCQNTYLVVPCFFYLACGPFYLYSRH 60
QY 61 DRGTYQWPLNKTATLGLFWTCVADLFYSFMRERSGFLAPVPLVSTLTGITTLLA 120
61 DRGTYQWPLNKTATLGLFWTCVADLFYSFMRERSGFLAPVPLVSTLTGITTLLA 120
Db 61 DRGTYQWPLNKTATLGLFWTCVADLFYSFMRERSGFLAPVPLVSTLTGITTLLA 120
QY 121 TPLIQLERRKGVSSGIMLTFTWLVALVCAIALIRSKIMTALKEDAQYDLFRDIFYFYS 180
121 TPLIQLERRKGVSSGIMLTFTWLVALVCAIALIRSKIMTALKEDAQYDLFRDIFYFYS 180
Db 121 TPLIQLERRKGVSSGIMLTFTWLVALVCAIALIRSKIMTALKEDAQYDLFRDIFYFYS 180
QY 181 LLLIQLVLSGFSRSPPLFSSTIHDNPECSSASFLSRTTFWWTGLIVRGYQPLEGSD 240
181 LLLIQLVLSGFSRSPPLFSSTIHDNPECSSASFLSRTTFWWTGLIVRGYQPLEGSD 240
Db 181 LLLIQLVLSGFSRSPPLFSSTIHDNPECSSASFLSRTTFWWTGLIVRGYQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRQPKVYSSDPAQPKSSKVDAN---EE 296
241 LMSLNKEDTSEQVVPVLVKNWKECAKTRQPKVYSSDPAQPKSSKVDAN---EE 296
Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRQPKVYSSDPAQPKSSKVDAN---EE 296
QY 297 VEAALIVKSPKEMNPSSLFKVLYKTFGYPFLMSFFFKAIHDLMMFSGQIILKILFYNDT 356
297 VEAALIVKSPKEMNPSSLFKVLYKTFGYPFLMSFFFKAIHDLMMFSGQIILKILFYNDT 356
Db 297 VEAALIVKSPKEMNPSSLFKVLYKTFGYPFLMSFFFKAIHDLMMFSGQIILKILFYNDT 356
QY 357 KAPDWQGYFYTVLLFVYACQLTILVHQYFHCIVSGMRKRTAVYGAVERALVYTSARK 416
357 KAPDWQGYFYTVLLFVYACQLTILVHQYFHCIVSGMRKRTAVYGAVERALVYTSARK 416
Db 357 KAPDWQGYFYTVLLFVYACQLTILVHQYFHCIVSGMRKRTAVYGAVERALVYTSARK 416
QY 356 SAPNMGQFYTGILFVACQILTLHQYFHCIVSGMRKRTAVYGAVERALVYTSARK 415
356 SAPNMGQFYTGILFVACQILTLHQYFHCIVSGMRKRTAVYGAVERALVYTSARK 415
Db 356 SAPNMGQFYTGILFVACQILTLHQYFHCIVSGMRKRTAVYGAVERALVYTSARK 415
QY 417 SSTVGEIVNLMSTDAQRFMDLATYIMNWSAPQVILALVILMLNIGPSTIAGYAVNWVM 476
417 SSTVGEIVNLMSTDAQRFMDLATYIMNWSAPQVILALVILMLNIGPSTIAGYAVNWVM 476
Db 417 SSTVGEIVNLMSTDAQRFMDLATYIMNWSAPQVILALVILMLNIGPSTIAGYAVNWVM 476
QY 416 TSTVGEIVNLMSTDAQRFMDLATYIMNWSAPQVILALVILMLNIGPSTIAGYAVNW 475
416 TSTVGEIVNLMSTDAQRFMDLATYIMNWSAPQVILALVILMLNIGPSTIAGYAVNW 475
Db 416 TSTVGEIVNLMSTDAQRFMDLATYIMNWSAPQVILALVILMLNIGPSTIAGYAVNW 475
QY 477 VPVNAVWAMKTKTYOYVAHMKSKDNRIKLMEILNGIKVLGLYAMELAFKQVLAIRBEL 536
477 VPVNAVWAMKTKTYOYVAHMKSKDNRIKLMEILNGIKVLGLYAMELAFKQVLAIRBEL 536
Db 477 VPVNAVWAMKTKTYOYVAHMKSKDNRIKLMEILNGIKVLGLYAMELAFKQVLAIRBEL 536
QY 476 VPIVNAVWAMKTKTYOYVAHMKSKDNRIKLMEILNGIKVLGLYAMELAFKQVLAIRBEL 535
476 VPIVNAVWAMKTKTYOYVAHMKSKDNRIKLMEILNGIKVLGLYAMELAFKQVLAIRBEL 535
Db 476 VPIVNAVWAMKTKTYOYVAHMKSKDNRIKLMEILNGIKVLGLYAMELAFKQVLAIRBEL 535
QY 537 KVLKSAVYLSAVGTFTWCTPPLVALCTFAVYVYTIENNTLDAQTAQVSLATFNILRPL 596
537 KVLKSAVYLSAVGTFTWCTPPLVALCTFAVYVYTIENNTLDAQTAQVSLATFNILRPL 596
Db 537 KVLKSAVYLSAVGTFTWCTPPLVALCTFAVYVYTIENNTLDAQTAQVSLATFNILRPL 596
QY 536 KVLKSAVYLSAVGTFTWCTPPLVALCTFAVYVYTIENNTLDAQTAQVSLATFNILRPL 595
536 KVLKSAVYLSAVGTFTWCTPPLVALCTFAVYVYTIENNTLDAQTAQVSLATFNILRPL 595
Db 536 KVLKSAVYLSAVGTFTWCTPPLVALCTFAVYVYTIENNTLDAQTAQVSLATFNILRPL 595
QY 597 NIIIPWYISSIVQASVSLKRLRITLSHELEPDSIERRPVKDGGCTNSITVRANTFTWARS 656
597 NIIIPWYISSIVQASVSLKRLRITLSHELEPDSIERRPVKDGGCTNSITVRANTFTWARS 656
Db 597 NIIIPWYISSIVQASVSLKRLRITLSHELEPDSIERRPVKDGGCTNSITVRANTFTWARS 656
QY 596 NIIIPWYISSIVQASVSLKRLRITLSHELEPDSIERRPVKDGGCTNSITVRANTFTWARS 653
596 NIIIPWYISSIVQASVSLKRLRITLSHELEPDSIERRPVKDGGCTNSITVRANTFTWARS 653
Db 596 NIIIPWYISSIVQASVSLKRLRITLSHELEPDSIERRPVKDGGCTNSITVRANTFTWARS 653
QY 657 DPEPLNGITPISIEGALVNAVGVGGKSSLSALAEADKVGHAIGSVVYVPOQAV 716
657 DPEPLNGITPISIEGALVNAVGVGGKSSLSALAEADKVGHAIGSVVYVPOQAV 716
Db 657 DPEPLNGITPISIEGALVNAVGVGGKSSLSALAEADKVGHAIGSVVYVPOQAV 716
QY 654 DPEPLNGITPISIEGALVNAVGVGGKSSLSALAEADKVGHAIGSVVYVPOQAV 713
654 DPEPLNGITPISIEGALVNAVGVGGKSSLSALAEADKVGHAIGSVVYVPOQAV 713
Db 654 DPEPLNGITPISIEGALVNAVGVGGKSSLSALAEADKVGHAIGSVVYVPOQAV 713
QY 717 IQNDSURENTLFCQLEBPYRSVIOACALLPDLRLTSGDRTEIEGKGNVLSGGQKORV 776
717 IQNDSURENTLFCQLEBPYRSVIOACALLPDLRLTSGDRTEIEGKGNVLSGGQKORV 776
Db 717 IQNDSURENTLFCQLEBPYRSVIOACALLPDLRLTSGDRTEIEGKGNVLSGGQKORV 776
QY 714 IQNATLEDNITPERENMESRYKRVIEACALLPDLRLTSGDRTEIEGKGNVLSGGQKORV 773
714 IQNATLEDNITPERENMESRYKRVIEACALLPDLRLTSGDRTEIEGKGNVLSGGQKORV 773
Db 714 IQNATLEDNITPERENMESRYKRVIEACALLPDLRLTSGDRTEIEGKGNVLSGGQKORV 773
QY 777 SLARAVYSNADYILFPDDPLSANDAHVGHKIFENVIGPKMLKXKTRILVTHSNSTYLPQVD 836
777 SLARAVYSNADYILFPDDPLSANDAHVGHKIFENVIGPKMLKXKTRILVTHSNSTYLPQVD 836
Db 777 SLARAVYSNADYILFPDDPLSANDAHVGHKIFENVIGPKMLKXKTRILVTHSNSTYLPQVD 836
QY 774 SLARAVYCNADYILFPDDPLSANDAHVGHKIFENVIGPKMLKXKTRILVTHSNSTYLPQVD 833
774 SLARAVYCNADYILFPDDPLSANDAHVGHKIFENVIGPKMLKXKTRILVTHSNSTYLPQVD 833
Db 774 SLARAVYCNADYILFPDDPLSANDAHVGHKIFENVIGPKMLKXKTRILVTHSNSTYLPQVD 833
QY 837 VITVWGGKISGEMSYOELLARPGAFAPFLRTAATYBOBDAEAGCVTVGSGP-GKEAKQ 895
837 VITVWGGKISGEMSYOELLARPGAFAPFLRTAATYBOBDAEAGCVTVGSGP-GKEAKQ 895
Db 837 VITVWGGKISGEMSYOELLARPGAFAPFLRTAATYBOBDAEAGCVTVGSGP-GKEAKQ 895
QY 834 TILVWTDGERISEMGSYOELLARPGAFAPFLRTAATYBOBDAEAGCVTVGSGP-GKEAKQ 888
834 TILVWTDGERISEMGSYOELLARPGAFAPFLRTAATYBOBDAEAGCVTVGSGP-GKEAKQ 888
Db 834 TILVWTDGERISEMGSYOELLARPGAFAPFLRTAATYBOBDAEAGCVTVGSGP-GKEAKQ 888
QY 896 MENGLVTDGAKQOLQRLSSSSSSYSGDI--SHHNSYALQLQRAKKEBTWTLMADKA 953
896 MENGLVTDGAKQOLQRLSSSSSSYSGDI--SHHNSYALQLQRAKKEBTWTLMADKA 953
Db 896 MENGLVTDGAKQOLQRLSSSSSSYSGDI--SHHNSYALQLQRAKKEBTWTLMADKA 953
QY 889 VENGVLVNDATGKLMRQLSNSSTYRGTEKSHQSHSTAELOQPLAEK-NSWKLTEADTA 947
889 VENGVLVNDATGKLMRQLSNSSTYRGTEKSHQSHSTAELOQPLAEK-NSWKLTEADTA 947
Db 889 VENGVLVNDATGKLMRQLSNSSTYRGTEKSHQSHSTAELOQPLAEK-NSWKLTEADTA 947
QY 954 QTOGVVLSTYVWDKMAIGLFIPLSLFLPNCNIVSALASNYMLSTWTDPIVNGTOEHTK 1013
954 QTOGVVLSTYVWDKMAIGLFIPLSLFLPNCNIVSALASNYMLSTWTDPIVNGTOEHTK 1013
Db 954 QTOGVVLSTYVWDKMAIGLFIPLSLFLPNCNIVSALASNYMLSTWTDPIVNGTOEHTK 1013

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:25 ; Search time 39.0811 Seconds
(without alignments)
4076.547 Million cell updates/sec

Title: US-10-665-283-6

Perfect score: 9901

Sequence: 1 MALRGCASAGSDPIMDMNV.....ASVAVAKRKPSPDLS 1927

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Issued Patents AA: *
2: /cgn2_6/ptodata/1/iaa/5 COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6 COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/H_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	79.4	1531	1 US-08-463-092B-4	Sequence 4, Appl1
2	7860	79.4	1531	1 US-08-462-109A-4	Sequence 4, Appl1
3	7860	79.4	1531	1 US-08-460-907B-4	Sequence 4, Appl1
4	7860	79.4	1531	2 US-08-463-179A-4	Sequence 4, Appl1
5	7860	79.4	1531	2 US-08-461-384B-4	Sequence 4, Appl1
6	7860	79.4	1531	2 US-09-647-140B-19	Sequence 19, Appl1
7	7849	79.3	1531	1 US-08-141-893-2	Sequence 2, Appl1
8	7849	79.3	1531	1 US-08-463-092B-2	Sequence 2, Appl1
9	7849	79.3	1531	1 US-08-462-109A-2	Sequence 2, Appl1
10	7849	79.3	1531	1 US-08-460-907B-2	Sequence 2, Appl1
11	7849	79.3	1531	2 US-08-463-179A-2	Sequence 2, Appl1
12	7849	79.3	1531	2 US-08-461-384B-2	Sequence 2, Appl1
13	7849	79.3	1531	2 US-08-407-207A-2	Sequence 2, Appl1
14	7002.5	70.7	1528	1 US-08-463-092B-6	Sequence 6, Appl1
15	7002.5	70.7	1528	1 US-08-462-109A-6	Sequence 6, Appl1
16	7002.5	70.7	1528	1 US-08-460-907B-6	Sequence 6, Appl1
17	7002.5	70.7	1528	2 US-08-463-179A-6	Sequence 6, Appl1
18	7002.5	70.7	1528	2 US-08-461-384B-6	Sequence 6, Appl1
19	4487.5	45.3	1527	2 US-09-647-140B-6	Sequence 6, Appl1
20	4487.5	45.3	1530	2 US-09-647-140B-33	Sequence 33, Appl1
21	3391.5	34.3	1503	2 US-09-647-140B-8	Sequence 8, Appl1
22	3385.5	34.2	1503	2 US-09-792-616-3	Sequence 3, Appl1
23	3218.5	32.5	1498	2 US-09-792-616-9	Sequence 3, Appl1
24	2374.5	24.0	1621	2 US-08-972-927-3	Sequence 3, Appl1
25	2349.5	23.7	1622	2 US-08-972-927-6	Sequence 6, Appl1
26	2266.5	22.9	1325	2 US-09-647-140B-2	Sequence 2, Appl1
27	2212.5	22.3	1464	2 US-10-012-896-1008	Sequence 1008, Ap

ALIGNMENTS

28	2195.5	22.2	1261	2 US-09-439-313-538	Sequence 538, App
29	2195.5	22.2	1261	2 US-09-636-215-538	Sequence 538, App
30	2195.5	22.2	1261	2 US-09-685-166A-538	Sequence 538, App
31	2195.5	22.2	1261	2 US-09-679-426-538	Sequence 538, App
32	2195.5	22.2	1261	2 US-09-759-143-538	Sequence 538, App
33	2195.5	22.2	1261	2 US-09-651-236-538	Sequence 538, App
34	2195.5	22.2	1261	2 US-09-657-279-538	Sequence 538, App
35	2195.5	22.2	1261	2 US-10-012-896-538	Sequence 538, App
36	2195.5	22.2	1261	2 US-10-012-896-1009	Sequence 1009, App
37	2118	21.4	1581	2 US-08-726-320-3	Sequence 3, Appl1
38	2118	21.4	1581	2 US-09-208-716-3	Sequence 3, Appl1
39	2115	21.4	1228	2 US-09-439-313-537	Sequence 537, App
40	2115	21.4	1228	2 US-09-636-215-537	Sequence 537, App
41	2115	21.4	1228	2 US-09-685-166A-537	Sequence 537, App
42	2115	21.4	1228	2 US-09-679-426-537	Sequence 537, App
43	2115	21.4	1228	2 US-09-759-143-537	Sequence 537, App
44	2115	21.4	1228	2 US-09-651-236-537	Sequence 537, App
45	2115	21.4	1228	2 US-09-657-279-537	Sequence 537, App

RESULT 1
US-08-463-092B-4
Sequence 4, Application US/08463092B
Patent No. 5766880
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,359
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-0928-4

Query Match 79.4%; Score 7860; DB 1; length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALRGFCSADGSDPLMNMNTWNTSNPFTKCFQNTVLVWPCRYLWACFFPYLYLSRH 60
DB 1 MALRGFCSADGSDPLMNMNTWNTSNPFTKCFQNTVLVWPCRYLWACFFPYLYLSRH 60
QY 61 DRGYIOMTPLNKTKTALGFLLMIYCMADLFYSFMRSGRIFLAPYFLVSPILLGITTLA 120
DB 61 DRGYIOMTPLNKTKTALGFLLMIYCMADLFYSFMRSGRIFLAPYFLVSPILLGITTLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMVALVYCALALRSKIMTALKEDAQVDLFRDITFYVYS 180
DB 121 TFLIQLERRKGVSSGIMLTFMVALVYCALALRSKIMTALKEDAQVDLFRDITFYVYS 180
QY 181 LLLIQLVLSGSDSPFSETIHDNPPCRESSASFLSITFWMTGLIVRGYRQPLBGSD 240
DB 181 LLLIQLVLSGSDSPFSETIHDNPPCRESSASFLSITFWMTGLIVRGYRQPLBGSD 240
QY 241 LMSLNKEDTSBOVVPVLVKNMKKCAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSBOVVPVLVKNMKKCAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPKEMNBSLFKVLYKTEGPFYFLMSFFKAIHDLMMFSGPOLIKLLIFVNDYKAPD 360
DB 301 IVKSPKEMNBSLFKVLYKTEGPFYFLMSFFKAIHDLMMFSGPOLIKLLIFVNDYKAPD 360
QY 361 WQGFYVYVLFVTAQCLQVLVHOYFHICFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYVYVLFVTAQCLQVLVHOYFHICFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDQRFMDLATYINMIWSAPLOYILALYMLNLGSPVLGAVAVMLMVVN 480
DB 421 GEIYNLMSVDQRFMDLATYINMIWSAPLOYILALYMLNLGSPVLGAVAVMLMVVN 480
QY 481 AYMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLKYAMELAFKQVLAIROBELKYLK 540
DB 481 AYMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLKYAMELAFKQVLAIROBELKYLK 540
QY 541 KSAVLASVGTFTWCTPFLVALCTPAVYVITDENNILDQAOTAFVSLAFNLRPLNLTLP 600
DB 541 KSAVLASVGTFTWCTPFLVALCTPAVYVITDENNILDQAOTAFVSLAFNLRPLNLTLP 600
QY 601 MVISIYOASVSLKRLRIFLSHEBELPDSIRRPVKDGGNSITVRATFTWARSDDPT 660
DB 601 MVISIYOASVSLKRLRIFLSHEBELPDSIRRPVKDGGNSITVRATFTWARSDDPT 660
QY 661 LMGITFSPGALVAVVGVQVCGKSSLLSALLAEMDKYEGHVALIKGSVAVYVPOQAMIOND 720
DB 661 LMGITFSPGALVAVVGVQVCGKSSLLSALLAEMDKYEGHVALIKGSVAVYVPOQAMIOND 720
QY 721 SLRENILFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKNVNSGQOKORVSLAR 780
DB 721 SLRENILFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKNVNSGQOKORVSLAR 780
QY 781 AVYSNADITYLPDDPLSAVDAAVGHKIPENVIQPKMKLNKRIILVTHSMSTLPQVDVITV 840
DB 781 AVYSNADITYLPDDPLSAVDAAVGHKIPENVIQPKMKLNKRIILVTHSMSTLPQVDVITV 840
QY 841 MSGGKISMSGYOBELADGAFAEFLRTYASTBOQDAEENGVTGVSFGKEAKOMENGM 900
DB 841 MSGGKISMSGYOBELADGAFAEFLRTYASTBOQDAEENGVTGVSFGKEAKOMENGM 900
QY 901 LVTDSAGKOLOROLSSSSSYSGDISRRHNSIAELQKAEKKEETWKLMEADKAOTGOYVL 960
DB 901 LVTDSAGKOLOROLSSSSSYSGDISRRHNSIAELQKAEKKEETWKLMEADKAOTGOYVL 960
```

```
DB 901 LVTDSAGKOLOROLSSSSSYSGDISRRHNSIAELQKAEKKEETWKLMEADKAOTGOYVL 960
QY 961 SVYWDYMKAIQGLFISFISIFLFMCNHYSAASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYMKAIQGLFISFISIFLFMCNHYSAASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQGIAVFGYSMAVSIIGIILASRCLHYDLHLSIIRSPMSFPERTPSGNLVNRFSKEL 1080
DB 1021 ALGISQGIAVFGYSMAVSIIGIILASRCLHYDLHLSIIRSPMSFPERTPSGNLVNRFSKEL 1080
QY 1081 DTYDSMTPEYIKMWSGLFNVIGACIVILATPIAIIIPPLGIYFVQRFYVASSROL 1140
DB 1081 DTYDSMTPEYIKMWSGLFNVIGACIVILATPIAIIIPPLGIYFVQRFYVASSROL 1140
QY 1141 KRLESVRSPPYSHFNFTLIGVSVIRAFEEQERFIHOSDKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVRSPPYSHFNFTLIGVSVIRAFEEQERFIHOSDKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLFCVNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLFCVNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMIOETAPSSMPOYGRVFERVCLRYREDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPMIOETAPSSMPOYGRVFERVCLRYREDLPVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSLTLGLFRINESAEGEIIIDGINIAKIGLHDLFKITIIIPQDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTLGLFRINESAEGEIIIDGINIAKIGLHDLFKITIIIPQDPVLF 1380
QY 1381 SSGIRANLDPFSQYSDIEWTSLFLAHLKDFVSLPKLDHECAGEGNSISVGOQOLVCL 1440
DB 1381 SSGIRANLDPFSQYSDIEWTSLFLAHLKDFVSLPKLDHECAGEGNSISVGOQOLVCL 1440
QY 1441 ARALLRKTILYIDEAFAVLDLETDLIQSTIRQFEDCTVLTAAHRLNTMDTRYIVL 1500
DB 1441 ARALLRKTILYIDEAFAVLDLETDLIQSTIRQFEDCTVLTAAHRLNTMDTRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531

RESULT 2
US-08-462-109A-4
Sequence 4, Application US/08462109A
Patent No. 5882875
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,109A
FILING DATE:
CLASSIFICATION: 435
APPLICATION NUMBER: 07/966,923
PRIOR APPLICATION DATA:
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
```

FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-4

Query Match 79.4%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDANVTMTNSNPDFTKCPONTVLVWVPCFYLMACPFYFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDANVTMTNSNPDFTKCPONTVLVWVPCFYLMACPFYFLYLSRH 60
QY 61 DRGVIQMTPLNKTKTALGFLIMIVCMADLFYSFWEBSRGIFLAFVFLVSPITLIGITLLA 120
DB 61 DRGVIQMTPLNKTKTALGFLIMIVCMADLFYSFWEBSRGIFLAFVFLVSPITLIGITLLA 120
QY 121 TPLIQERRKGVSSGIMLTFMLVALCALAIIKSKIMTAKEDAOVDLPDITFYFYS 180
DB 121 TPLIQERRKGVSSGIMLTFMLVALCALAIIKSKIMTAKEDAOVDLPDITFYFYS 180
QY 121 TPLIQERRKGVSSGIMLTFMLVALCALAIIKSKIMTAKEDAOVDLPDITFYFYS 180
DB 121 TPLIQERRKGVSSGIMLTFMLVALCALAIIKSKIMTAKEDAOVDLPDITFYFYS 180
QY 181 LLLIQVLSCGSDSPFSETIHDNCPRESSASFLSRITPMTTGLIVGYRQPLGSD 240
DB 181 LLLIQVLSCGSDSPFSETIHDNCPRESSASFLSRITPMTTGLIVGYRQPLGSD 240
QY 181 LLLIQVLSCGSDSPFSETIHDNCPRESSASFLSRITPMTTGLIVGYRQPLGSD 240
DB 181 LLLIQVLSCGSDSPFSETIHDNCPRESSASFLSRITPMTTGLIVGYRQPLGSD 240
QY 241 LMSLNKEDTSBOVPLVVKWKKCACTRKQPVKVVYSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSBOVPLVVKWKKCACTRKQPVKVVYSKDPAPKESKVDANEVEAL 300
QY 241 LMSLNKEDTSBOVPLVVKWKKCACTRKQPVKVVYSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSBOVPLVVKWKKCACTRKQPVKVVYSKDPAPKESKVDANEVEAL 300
QY 301 IVKSPQKKNPSPFKVITKTPGPFLMSFPFKAIHDLMTSGPQILKLIKFNVDTPAD 360
DB 301 IVKSPQKKNPSPFKVITKTPGPFLMSFPFKAIHDLMTSGPQILKLIKFNVDTPAD 360
QY 301 IVKSPQKKNPSPFKVITKTPGPFLMSFPFKAIHDLMTSGPQILKLIKFNVDTPAD 360
DB 301 IVKSPQKKNPSPFKVITKTPGPFLMSFPFKAIHDLMTSGPQILKLIKFNVDTPAD 360
QY 361 MOGFYTVLLFVTRACLOTVLVHOVFIHCVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
DB 361 MOGFYTVLLFVTRACLOTVLVHOVFIHCVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
QY 361 MOGFYTVLLFVTRACLOTVLVHOVFIHCVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
DB 361 MOGFYTVLLFVTRACLOTVLVHOVFIHCVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAOREMDLATYINMIWSAPLOVITLALYLLMLNIGPSVLGAVAVMLAMPVN 480
DB 421 GEIVNLMSVDAOREMDLATYINMIWSAPLOVITLALYLLMLNIGPSVLGAVAVMLAMPVN 480
QY 421 GEIVNLMSVDAOREMDLATYINMIWSAPLOVITLALYLLMLNIGPSVLGAVAVMLAMPVN 480
DB 421 GEIVNLMSVDAOREMDLATYINMIWSAPLOVITLALYLLMLNIGPSVLGAVAVMLAMPVN 480
QY 481 AVNMAKTKTYQVAHMSKDNRIKIMNEILNGIKVLYAMELAFKDVLAIRQELKVLK 540
DB 481 AVNMAKTKTYQVAHMSKDNRIKIMNEILNGIKVLYAMELAFKDVLAIRQELKVLK 540
QY 481 AVNMAKTKTYQVAHMSKDNRIKIMNEILNGIKVLYAMELAFKDVLAIRQELKVLK 540
DB 481 AVNMAKTKTYQVAHMSKDNRIKIMNEILNGIKVLYAMELAFKDVLAIRQELKVLK 540
QY 541 KSAVLSAVGTFTWCTPELVALCTFAVYVITDENNIIIDAQTAFLALFNILRPILP 600
DB 541 KSAVLSAVGTFTWCTPELVALCTFAVYVITDENNIIIDAQTAFLALFNILRPILP 600
QY 541 KSAVLSAVGTFTWCTPELVALCTFAVYVITDENNIIIDAQTAFLALFNILRPILP 600
DB 541 KSAVLSAVGTFTWCTPELVALCTFAVYVITDENNIIIDAQTAFLALFNILRPILP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKGGGNTSITVNAATFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKGGGNTSITVNAATFTWARSDDPT 660
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKGGGNTSITVNAATFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKGGGNTSITVNAATFTWARSDDPT 660
QY 661 LNTGTFSPREGALVAVVGVGGCGSSLSLALAMDMVEGHVAKGSVAAYVPOAMQND 720
DB 661 LNTGTFSPREGALVAVVGVGGCGSSLSLALAMDMVEGHVAKGSVAAYVPOAMQND 720
QY 661 LNTGTFSPREGALVAVVGVGGCGSSLSLALAMDMVEGHVAKGSVAAYVPOAMQND 720
DB 661 LNTGTFSPREGALVAVVGVGGCGSSLSLALAMDMVEGHVAKGSVAAYVPOAMQND 720
QY 721 SLRENILFGQLEBPYRSVIQACALLPDEIILPSGDRTEIGEGVNLSSGQKQVSLAR 780

DB 721 SLRENILFGQLEBPYRSVIQACALLPDEIILPSGDRTEIGEGVNLSSGQKQVSLAR 780
QY 781 AVYSNADIYFDDBLSAVDAHVGKHIFENVIIGPKMKTKTRILVTHSMGILPOVDVYIV 840
DB 781 AVYSNADIYFDDBLSAVDAHVGKHIFENVIIGPKMKTKTRILVTHSMGILPOVDVYIV 840
QY 841 MSGKISEMSYQELARDGAPAEFLRTVSTOEOAEENGTVGSGPKKAKOMNGM 900
DB 841 MSGKISEMSYQELARDGAPAEFLRTVSTOEOAEENGTVGSGPKKAKOMNGM 900
QY 841 MSGKISEMSYQELARDGAPAEFLRTVSTOEOAEENGTVGSGPKKAKOMNGM 900
DB 841 MSGKISEMSYQELARDGAPAEFLRTVSTOEOAEENGTVGSGPKKAKOMNGM 900
QY 901 LVYDSAGKQORQLSSSSSYSGDISRRHNSTAELOKAEKETKLMKADKQOTGVKL 960
DB 901 LVYDSAGKQORQLSSSSSYSGDISRRHNSTAELOKAEKETKLMKADKQOTGVKL 960
QY 901 LVYDSAGKQORQLSSSSSYSGDISRRHNSTAELOKAEKETKLMKADKQOTGVKL 960
DB 901 LVYDSAGKQORQLSSSSSYSGDISRRHNSTAELOKAEKETKLMKADKQOTGVKL 960
QY 961 SVYWDYKKAIGLISFLSIFLPMCNHVSALASNYWSLMTDDPIVNGTOHTTKVRLSVYG 1020
DB 961 SVYWDYKKAIGLISFLSIFLPMCNHVSALASNYWSLMTDDPIVNGTOHTTKVRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSGIILASRCLHYDLHSILRSPMSFEERTPSGNLVNRSKEL 1080
DB 1021 ALGISOGIAVFGYSMAVSGIILASRCLHYDLHSILRSPMSFEERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPVYIKMFMSLFVNIAGCTVILATPILAIIPPLGIYFFVORFYAASSROL 1140
DB 1081 DTVDSMIPVYIKMFMSLFVNIAGCTVILATPILAIIPPLGIYFFVORFYAASSROL 1140
QY 1141 KRLSVRSAPVYSHFNETLLGVSIVRAFEQBERFIHQSDLKVDENOKAYYPSIVANWLA 1200
DB 1141 KRLSVRSAPVYSHFNETLLGVSIVRAFEQBERFIHQSDLKVDENOKAYYPSIVANWLA 1200
QY 1141 KRLSVRSAPVYSHFNETLLGVSIVRAFEQBERFIHQSDLKVDENOKAYYPSIVANWLA 1200
DB 1141 KRLSVRSAPVYSHFNETLLGVSIVRAFEQBERFIHQSDLKVDENOKAYYPSIVANWLA 1200
QY 1201 VRLCVCNCIVLPAALPAVYSRHSLSAGLVGSVSVSLQVTTIYLMVLRMSSEMETIVA 1260
DB 1201 VRLCVCNCIVLPAALPAVYSRHSLSAGLVGSVSVSLQVTTIYLMVLRMSSEMETIVA 1260
QY 1261 VERLKEVSETEKAPMOIQETAPSSWPQVGRVEFRNYCGRYEDLDVLRHINVTNGG 1320
DB 1261 VERLKEVSETEKAPMOIQETAPSSWPQVGRVEFRNYCGRYEDLDVLRHINVTNGG 1320
QY 1321 EKVGIWGTAGKSSLTGLFRINESAGEI11DGINIAKIGLHDLRFKTTI1PODEVLF 1380
DB 1321 EKVGIWGTAGKSSLTGLFRINESAGEI11DGINIAKIGLHDLRFKTTI1PODEVLF 1380
QY 1381 SGLRNLDPFQYSDBEVWTSLELAHKDFVSLPDKLDHECAEGGENISVQORQVCL 1440
DB 1381 SGLRNLDPFQYSDBEVWTSLELAHKDFVSLPDKLDHECAEGGENISVQORQVCL 1440
QY 1441 ARALLRKTKLIVDEATAVDLETTDLOSTIRTOFEDCTVLTIAHRLNTIMYTRYIVL 1500
DB 1441 ARALLRKTKLIVDEATAVDLETTDLOSTIRTOFEDCTVLTIAHRLNTIMYTRYIVL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531

RESULT 3
US-08-460-907B-4
Sequence 4, Application US/08460907B
Patent No. 5891724
GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
RESISTANCE ON A CELL
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-4

Query Match      79.4%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCAGSDSDPLMDNNTWNTNSPDFTKCFQNTVAVWVPCFYLMACFPFFYLYLSRH 60
DB 1 MALRGFCAGSDSDPLMDNNTWNTNSPDFTKCFQNTVAVWVPCFYLMACFPFFYLYLSRH 60
QY 61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPILLGITLLA 120
DB 61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPILLGITLLA 120
QY 121 TFLIQLERRKVOSSGIMLTFWLVALVCALAILRKSITALKEDQVDFRDIITYVFS 180
DB 121 TFLIQLERRKVOSSGIMLTFWLVALVCALAILRKSITALKEDQVDFRDIITYVFS 180
QY 121 TFLIQLERRKVOSSGIMLTFWLVALVCALAILRKSITALKEDQVDFRDIITYVFS 180
DB 121 TFLIQLERRKVOSSGIMLTFWLVALVCALAILRKSITALKEDQVDFRDIITYVFS 180
QY 181 LLLIQLVSCFSDSPLESETIHDPNCPBESSASFLSITTFWITGLIVRGYROPLEGGSD 240
DB 181 LLLIQLVSCFSDSPLESETIHDPNCPBESSASFLSITTFWITGLIVRGYROPLEGGSD 240
QY 241 LMSLNKEDTSEBOVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPQKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEBOVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPQKSSKVDANEVEAL 300
QY 301 IVKBPQKWNNSLPFVLYKTGPFYFLMSFPFKALHDLMMFSGPOLIKLILKRVNDTKAPD 360
DB 301 IVKBPQKWNNSLPFVLYKTGPFYFLMSFPFKALHDLMMFSGPOLIKLILKRVNDTKAPD 360
QY 361 WQGFYTYLTFYTLACLOTLYLHOYFHCIFVSGMRKIKTAVIGAVYKALVITNSAKKSSV 420
DB 361 WQGFYTYLTFYTLACLOTLYLHOYFHCIFVSGMRKIKTAVIGAVYKALVITNSAKKSSV 420
QY 421 GEIVNLSVDAQRFMDLATYINIMWSAPLOVTLALYLLMLNLGSPVLGVAVMVLAVPN 480
```

```

|||||
421 GEIVNLSVDAQRFMDLATYINIMWSAPLOVTLALYLLMLNLGSPVLGVAVMVLAVPN 480
QY AVAMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLKLYAMELAFKDYALIROBELKYLK 540
DB AVAMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLKLYAMELAFKDYALIROBELKYLK 540
QY 481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLKLYAMELAFKDYALIROBELKYLK 540
DB 481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLKLYAMELAFKDYALIROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNNILDAQTAVYSALPNILRFLPNILP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNNILDAQTAVYSALPNILRFLPNILP 600
QY 601 MVISSTIQAASVSLKRLAIFLSHELEPDSIERRVKGGCGNSTTNNATFTMARSDPPT 660
DB 601 MVISSTIQAASVSLKRLAIFLSHELEPDSIERRVKGGCGNSTTNNATFTMARSDPPT 660
QY 661 LINGITFIPEGALVAVVGVQCGKSLSLSALLAEMDYEGHVAIKGSVAAYVPOQAMQND 720
DB 661 LINGITFIPEGALVAVVGVQCGKSLSLSALLAEMDYEGHVAIKGSVAAYVPOQAMQND 720
QY 721 SLRENILFGCQLEBPYRSVYIQAACALLPDLLEILPSGURTEIGEKVNILSGGQKORVSLAR 780
DB 721 SLRENILFGCQLEBPYRSVYIQAACALLPDLLEILPSGURTEIGEKVNILSGGQKORVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDHVGKHIPEVNYGPKGMLKNKTRILVTHSMSTLPQVDVTV 840
DB 781 AVYSNADIYLFDDPLSAVDHVGKHIPEVNYGPKGMLKNKTRILVTHSMSTLPQVDVTV 840
QY 841 MSGGKISEMSYQELRLARDGAFAEFLRTYASTEOQDAEENGVTGVSQPKAEALQEMNGM 900
DB 841 MSGGKISEMSYQELRLARDGAFAEFLRTYASTEOQDAEENGVTGVSQPKAEALQEMNGM 900
QY 901 LVYDSAGKQLOROLSSSSSYSGDISRRHNSTAELQKAEXKEETWKLMEADKAOTGOVKL 960
DB 901 LVYDSAGKQLOROLSSSSSYSGDISRRHNSTAELQKAEXKEETWKLMEADKAOTGOVKL 960
QY 961 SVYWDYMKAIQLFISPLSIFLMCNHVSALASNWLSLMTDDPIVNGQHEHTKRLSYVG 1020
DB 961 SVYWDYMKAIQLFISPLSIFLMCNHVSALASNWLSLMTDDPIVNGQHEHTKRLSYVG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCIHDVLLHSILRSPMSFERPFGNULNRFSEL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCIHDVLLHSILRSPMSFERPFGNULNRFSEL 1080
QY 1081 DTVDSMIPEVIKMGMSLFVNIGACIVILATPIAIIIPPLGIYFFVQGFYVASSROL 1140
DB 1081 DTVDSMIPEVIKMGMSLFVNIGACIVILATPIAIIIPPLGIYFFVQGFYVASSROL 1140
QY 1141 KRLSVSRSPYSHFNETLLGVSYITRAPEQERTIHQSDLKVDENQKAYYSIYANRWLA 1200
DB 1141 KRLSVSRSPYSHFNETLLGVSYITRAPEQERTIHQSDLKVDENQKAYYSIYANRWLA 1200
QY 1201 VRLSCVGNCTVLPALFAVVISRHSLSAGLVGLSVSYLQVTTYVNLVYRMSSEMETNIVA 1260
DB 1201 VRLSCVGNCTVLPALFAVVISRHSLSAGLVGLSVSYLQVTTYVNLVYRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSPSPQVGRVFFRNVCYLRBDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPSPSPQVGRVFFRNVCYLRBDLPVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAKGSLTGLFRINESABGEIITGGINIAKIGLHDLRKITITIPDDPVLF 1380
DB 1321 EKVGIVGRTGAKGSLTGLFRINESABGEIITGGINIAKIGLHDLRKITITIPDDPVLF 1380
QY 1381 SSGSLRMNLDPSQVSDSEVWTSLELAHLKDVVSALPDLDHCBAGGNSLVGQRQVLC 1440
DB 1381 SSGSLRMNLDPSQVSDSEVWTSLELAHLKDVVSALPDLDHCBAGGNSLVGQRQVLC 1440
QY 1441 ARALLRKTKIIVLDEATAAVDELFTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
DB 1441 ARALLRKTKIIVLDEATAAVDELFTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIOEGAPSDLIQQRGLFYSAKDAGLV 1531
```


Db 1501 DKGEIOEXGAPSDLLQORGLFVSMADQAGLV 1531

RESULT 4
US-08-463-179A-4
Sequence 4, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deconcti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-4

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFGSADGSDPLMDMNTWNTSNDFYKCFONTVLVWPCFYLMACFPFYLYSRH 60
Db 1 MALRGFGSADGSDPLMDMNTWNTSNDFYKCFONTVLVWPCFYLMACFPFYLYSRH 60
QY 61 DRGVIQMTPLNKTALGFLIMIVQWADLFYSPWERSGIFLAVPLVSPLLGITTLLA 120
Db 61 DRGVIQMTPLNKTALGFLIMIVQWADLFYSPWERSGIFLAVPLVSPLLGITTLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVVALYCALAILRSKIMTALKEDAQVDLFPDITFYVFS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVVALYCALAILRSKIMTALKEDAQVDLFPDITFYVFS 180
QY 181 LLLIQLVLSGFSDDSPFLSETIHDPNCPRESSASFLSRITFWMTGLIVRGYRPLRGSD 240
Db 181 LLLIQLVLSGFSDDSPFLSETIHDPNCPRESSASFLSRITFWMTGLIVRGYRPLRGSD 240
QY 241 LMSLNKEDTSRQVAVLVKMKWKECAKTRKQPVKVVYSSKDPAPKSSSKVDANBEVVAL 300

Db 241 LMSLNKEDTSRQVAVLVKMKWKECAKTRKQPVKVVYSSKDPAPKSSSKVDANBEVVAL 300
QY 301 IVKSPQKEMNPSLTKVLYKTFGRPYFLMSFFPKA1HDLMSGPOILKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLTKVLYKTFGRPYFLMSFFPKA1HDLMSGPOILKLLIKFVNDTKAPD 360
QY 361 WQGYFYVLLFEVTCLOTLVLVHQYFHI CFSGGMKIKTAIVGAVYRKALVTNSARKSTV 420
Db 361 WQGYFYVLLFEVTCLOTLVLVHQYFHI CFSGGMKIKTAIVGAVYRKALVTNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMWSAPLOVYIALYLLMLNLGSPVLGAVAVMLAMPVN 480
Db 421 GEIVNLMSVDAQRFMDLATYINMWSAPLOVYIALYLLMLNLGSPVLGAVAVMLAMPVN 480
QY 481 AVMAKTKTYOVAMHKSKDNRIKLMNEILNGIKVLYAMELAKDYLAIROBELKVLK 540
Db 481 AVMAKTKTYOVAMHKSKDNRIKLMNEILNGIKVLYAMELAKDYLAIROBELKVLK 540
QY 541 KSAVLAVGFTWCTPFLVALCTFAVYVITDENNILDQAOTAFVSLAFNLRPPLNLP 600
Db 541 KSAVLAVGFTWCTPFLVALCTFAVYVITDENNILDQAOTAFVSLAFNLRPPLNLP 600
QY 601 MVISIYQASVSLKRLIFLSHEELBPSDIERRPVKGGGNTSITVRNATFTMARSDPT 660
Db 601 MVISIYQASVSLKRLIFLSHEELBPSDIERRPVKGGGNTSITVRNATFTMARSDPT 660
QY 661 LMGITFSIPBGALVAVVGVGCGKSSLSALLAEMDKVEGHVAKGSVAVVPOQAMIQND 720
Db 661 LMGITFSIPBGALVAVVGVGCGKSSLSALLAEMDKVEGHVAKGSVAVVPOQAMIQND 720
QY 721 SLRENILFGCOLBEPYRYSYIQAALLPDLEILBPGRTEIGEGVNLSSGQKQVSLAR 780
Db 721 SLRENILFGCOLBEPYRYSYIQAALLPDLEILBPGRTEIGEGVNLSSGQKQVSLAR 780
QY 781 AVVSNADIYLPDDPLSAVDHVGKHIPENYIGPQMKKNTRIILVTHSMSTLPQVDYIIV 840
Db 781 AVVSNADIYLPDDPLSAVDHVGKHIPENYIGPQMKKNTRIILVTHSMSTLPQVDYIIV 840
QY 841 MSGGKISEMGSYOBLLARDGAFAEFLRTVASTBOBODABENGVTGSGPGKEAKOMENG 900
Db 841 MSGGKISEMGSYOBLLARDGAFAEFLRTVASTBOBODABENGVTGSGPGKEAKOMENG 900
QY 901 LVYDSAKQIQRQULSSSSSYSGDISRHNSYAELOKBAKKEETWKMEADKQOTGOYKL 960
Db 901 LVYDSAKQIQRQULSSSSSYSGDISRHNSYAELOKBAKKEETWKMEADKQOTGOYKL 960
QY 961 SVYWDYMKALGLFISFLSIFLFCMCHVSALASNYWLSLMTDDPIVNGTOEHTKYRLSYG 1020
Db 961 SVYWDYMKALGLFISFLSIFLFCMCHVSALASNYWLSLMTDDPIVNGTOEHTKYRLSYG 1020
QY 1021 ALGISOGIAVFGYMAVSIGAILASRCLAHVLDLHSILRSFMSFPERTPSGNLVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSIGAILASRCLAHVLDLHSILRSFMSFPERTPSGNLVNRFSKEL 1080
QY 1081 DTYDSMIPBYIKNFMGSLFNIVIGACIVILATPAAIIIPPLGLITFFVORFYVASSROL 1140
Db 1081 DTYDSMIPBYIKNFMGSLFNIVIGACIVILATPAAIIIPPLGLITFFVORFYVASSROL 1140
QY 1141 KRLSVSRSPVYSFNFETLLGVSIVIRAFEEQERFIHOSDLKVDENQAAVYPSIVANMFLA 1200
Db 1141 KRLSVSRSPVYSFNFETLLGVSIVIRAFEEQERFIHOSDLKVDENQAAVYPSIVANMFLA 1200
QY 1201 VRLCVCNCIVLFAALFAVLSRHSLSAGVLGSYSVLQVTTYLNMVLVRRSSSEMETNIVA 1260
Db 1201 VRLCVCNCIVLFAALFAVLSRHSLSAGVLGSYSVLQVTTYLNMVLVRRSSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPNOIOETAPSSWPQYGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPNOIOETAPSSWPQYGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVGIVGRTAGKSSSLTGLFRINESAEGEIIIDGIIAKIGLHDLRFKTTIIPQDVLVLF 1380

Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEI1IDGINIAKIGLHDLRFKTTITIPDDVLF 1380
Qy 1381 SGLSRMLDPPSOYSDDEEVTSLSLAHLKDFVSALPDLDBHCAEGENLSVGROLYCL 1440
Db 1381 SGLSRMLDPPSOYSDDEEVTSLSLAHLKDFVSALPDLDBHCAEGENLSVGROLYCL 1440
Qy 1441 ARALLRRTKIIVLDBATAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDTRVIVL 1500
Db 1441 ARALLRRTKIIVLDBATAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDTRVIVL 1500
Qy 1501 DKGEIOEGARSDLLQORGLFYSMKXGGLV 1531
Db 1501 DKGEIOEGARSDLLQORGLFYSMKXGGLV 1531

RESULT 5
US-08-461-384B-4
Sequence 4, Application US/08461384B
Patent No. 6025473
GENERAL INFORMATION:
APPLICANT: Coley, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,384B
FILING DATE: 05-JUN-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01547
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-384B-4

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCAGSDSDPLMDNNTVNTNSPDFTKCFONTVLVWVPCFYLMACFPFYFLYSRH 60
Db 1 MALRGFCAGSDSDPLMDNNTVNTNSPDFTKCFONTVLVWVPCFYLMACFPFYFLYSRH 60
Qy 61 DRGYIQMTPLNKTALGFLMIWCMADLFYSFWERSRGIFLAAPVFLVSPILLGITTLIA 120
Db 61 DRGYIQMTPLNKTALGFLMIWCMADLFYSFWERSRGIFLAAPVFLVSPILLGITTLIA 120

Db 61 DRGYIQMTPLNKTALGFLMIWCMADLFYSFWERSRGIFLAAPVFLVSPILLGITTLIA 120
Qy 121 TFLIQLERRKGVSSGIMLTFWLVALVCALA1LSRKIMTALKEDAQVDLFEDITFYVYFS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVALVCALA1LSRKIMTALKEDAQVDLFEDITFYVYFS 180
Qy 181 LLLIQLVLSGCSFDRSPLEFSETIHDNPNCPRESSASFLSRITFWMTGLIVRGYROPLEBSD 240
Db 181 LLLIQLVLSGCSFDRSPLEFSETIHDNPNCPRESSASFLSRITFWMTGLIVRGYROPLEBSD 240
Qy 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVVYVSSKQPAOPKSSKVDANEVVAL 300
Db 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVVYVSSKQPAOPKSSKVDANEVVAL 300
Qy 301 IVKSPKEMNSLPKVLVYKTFGPYFLMSFFPKAIHDLMMFSGPOLKLLIFVNDYKAPD 360
Db 301 IVKSPKEMNSLPKVLVYKTFGPYFLMSFFPKAIHDLMMFSGPOLKLLIFVNDYKAPD 360
Qy 361 WQGYFYTVLLFVYACLOTVLVHQYFHI CFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFYTVLLFVYACLOTVLVHQYFHI CFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420
Qy 421 GEIVNLSVDAQRFMDLATYINMIMSAPLOVITLALYLLMLNLSVLAGVAVMLMVPVN 480
Db 421 GEIVNLSVDAQRFMDLATYINMIMSAPLOVITLALYLLMLNLSVLAGVAVMLMVPVN 480
Qy 481 AVAMAKTKTYOAHMKSNDNRKIKLMBELNGLIKVLYKUYAMELAFKDYALROBELKYLK 540
Db 481 AVAMAKTKTYOAHMKSNDNRKIKLMBELNGLIKVLYKUYAMELAFKDYALROBELKYLK 540
Qy 541 KSAVLSAVGTTWCTPPLVALCTPVAVYTTIDENNIIIDAOCTAFVSLAFNLRPPNLTLP 600
Db 541 KSAVLSAVGTTWCTPPLVALCTPVAVYTTIDENNIIIDAOCTAFVSLAFNLRPPNLTLP 600
Qy 601 MVISIVASVSLRIRIFLSHEELBPSIERRPVKOGGNSITVRATFTMARSDPT 660
Db 601 MVISIVASVSLRIRIFLSHEELBPSIERRPVKOGGNSITVRATFTMARSDPT 660
Qy 661 LMGITFSPREGALVAVVQVCGKSSLSALLAEMDKVEGVAIKGSVAVYVPOQAMIOND 720
Db 661 LMGITFSPREGALVAVVQVCGKSSLSALLAEMDKVEGVAIKGSVAVYVPOQAMIOND 720
Qy 721 SLRENILFGCOLLEBPYRSVIOACALLPDLILPSGDRTEIGEKGVNLSGQOKORVSLAR 780
Db 721 SLRENILFGCOLLEBPYRSVIOACALLPDLILPSGDRTEIGEKGVNLSGQOKORVSLAR 780
Qy 781 AVYSNADIVLFPDDPLSANDAHVGHIFENVIGPKMLKNKRIILVTHSMSTYLPQVDVITV 840
Db 781 AVYSNADIVLFPDDPLSANDAHVGHIFENVIGPKMLKNKRIILVTHSMSTYLPQVDVITV 840
Qy 841 MSGGKI SEMGSYOEILARDGAFAEFLRTYASTEOQDAEENGVTGVSQPGKEAKOMENG 900
Db 841 MSGGKI SEMGSYOEILARDGAFAEFLRTYASTEOQDAEENGVTGVSQPGKEAKOMENG 900
Qy 901 LVYDSAGKOLORQLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADKAQTGYVL 960
Db 901 LVYDSAGKOLORQLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADKAQTGYVL 960
Qy 961 SVYWDYMKALGLFISFLSIFLPMCHVSA1SNWLSLMTDDPIVNGOHEHTKPVLSYVG 1020
Db 961 SVYWDYMKALGLFISFLSIFLPMCHVSA1SNWLSLMTDDPIVNGOHEHTKPVLSYVG 1020
Qy 1021 ALGISOGIAVFGYSMAVSIIGLILASRCHAVDLHSILSPMSFFERTSGNULVNFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGLILASRCHAVDLHSILSPMSFFERTSGNULVNFSEKL 1080
Qy 1081 DTVDSMIEPVIKMFGLSFNVIAGACTIVLLATPIAIIIPGLGIYFFVQRPYVASRQL 1140
Db 1081 DTVDSMIEPVIKMFGLSFNVIAGACTIVLLATPIAIIIPGLGIYFFVQRPYVASRQL 1140
Qy 1141 KRLSEVSRSPLYSHNETLLGVSVYTRAFEBQERPHOSDLYKDEQKAYYSIVANRWLA 1200
Db 1141 KRLSEVSRSPLYSHNETLLGVSVYTRAFEBQERPHOSDLYKDEQKAYYSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVIRSRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVIRSRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVEFRNYCLARYEDDLFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVEFRNYCLARYEDDLFVLRHINVTINGG 1320
QY 1321 EKVGIVRTGAGKSLTGLFRINESAGEIINGINAKIGLHDLRPKITIIPODPVLF 1380
Db 1321 EKVGIVRTGAGKSLTGLFRINESAGEIINGINAKIGLHDLRPKITIIPODPVLF 1380
QY 1381 SGSLRKMIDPPSOXSDEEWTSLFLAKHDKDFVSALPDKLDEHCEAGGENUSVGOROLVCL 1440
Db 1381 SGSLRKMIDPPSOXSDEEWTSLFLAKHDKDFVSALPDKLDEHCEAGGENUSVGOROLVCL 1440
QY 1441 ABALLRKTILVDEATAVADLETDDLIOSTIRTOFEDCTVLTAAHRLNTIMDYTRVIVL 1500
Db 1441 ABALLRKTILVDEATAVADLETDDLIOSTIRTOFEDCTVLTAAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOGYGAPSDLLQORGLFYSMADAGLV 1531
Db 1501 DKGEIOGYGAPSDLLQORGLFYSMADAGLV 1531

RESULT 6
US-09-647-140B-19
Sequence 19, Application US/09647140B
Patent No. 6803184
GENERAL INFORMATION:
APPLICANT: Fox Chase Cancer Center
APPLICANT: Kruth, Gary D.
APPLICANT: Lee, Kun
APPLICANT: Belinsky, Martin G.
APPLICANT: Bain, Lisa J.
TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
TITLE OF INVENTION: Nucleic Acids and Methods of Use Thereof
FILE REFERENCE: PCCC 98-02
CURRENT APPLICATION NUMBER: US/09/647,140B
CURRENT FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: PCT/US99/06644
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/079,759
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/095,153
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-647-140B-19

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMANTMTNTPDKCFONTLVVWPCFYLACFPFYFLYSRH 60
Db 1 MALRGFCSADGSDPLMDMANTMTNTPDKCFONTLVVWPCFYLACFPFYFLYSRH 60
QY 61 DRGYIOMTPIKNTKTALGFLIMIVCMADLPFYSFWERSRGIFLAPVFLVSTLGIITLLA 120
Db 61 DRGYIOMTPIKNTKTALGFLIMIVCMADLPFYSFWERSRGIFLAPVFLVSTLGIITLLA 120
QY 121 TPLIOLERRKGVSSGIMLFEMLVAYCALAIIISKIMTALKEBAQVDLRODTFFYYFS 180
Db 121 TPLIOLERRKGVSSGIMLFEMLVAYCALAIIISKIMTALKEBAQVDLRODTFFYYFS 180
QY 181 LLLIQLVLSCFSDRSPFLFSETHIDPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQLVLSCFSDRSPFLFSETHIDPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240

Db 181 LLLIQLVLSCFSDRSPFLFSETHIDPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSBOVVPVYVKNWKECACTRRQPKYVNVSSSDPAOPRESSVYDANEVEAL 300
Db 241 LMSLNKEDTSBOVVPVYVKNWKECACTRRQPKYVNVSSSDPAOPRESSVYDANEVEAL 300
QY 301 IVKSPQKEMPSSLFKVLKTFPGPYFLMSFFFKAIHDMMSGPOILLKLIKFNVDTAPD 360
Db 301 IVKSPQKEMPSSLFKVLKTFPGPYFLMSFFFKAIHDMMSGPOILLKLIKFNVDTAPD 360
QY 361 WQGYFTVTLFVTAQLOTLVLAHOYFHI CFVSGMRIKTAVI GAVYRKALVITNSARKSTV 420
Db 361 WQGYFTVTLFVTAQLOTLVLAHOYFHI CFVSGMRIKTAVI GAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAORFMDLATYINMISAPLOVITLALYLMNLGPSVLAGVAVWVLMVNV 480
Db 421 GEIVNLMSVDAORFMDLATYINMISAPLOVITLALYLMNLGPSVLAGVAVWVLMVNV 480
QY 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDKVLAIROBELKVLK 540
Db 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDKVLAIROBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNII DAOFAFVSLALFNILRFPNLILP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNII DAOFAFVSLALFNILRFPNLILP 600
QY 601 MVTSSIVQASVSLKRLIFLSHSELBPDSIERRPVKOGGTSITTENATFTMARSDPT 660
Db 601 MVTSSIVQASVSLKRLIFLSHSELBPDSIERRPVKOGGTSITTENATFTMARSDPT 660
QY 661 LMGITFSIPGALVAVVGOGCKSSLSLALBMDVBEHVAIKGSVAVPOQAWI QND 720
Db 661 LMGITFSIPGALVAVVGOGCKSSLSLALBMDVBEHVAIKGSVAVPOQAWI QND 720
QY 721 SLRENILFGQLEBPYRSVIOACALPDLEILPSCGRTEIGEXGVNLSSGQKORVSLAR 780
Db 721 SLRENILFGQLEBPYRSVIOACALPDLEILPSCGRTEIGEXGVNLSSGQKORVSLAR 780
QY 781 AVYSNADIVLPDPLSADVAHVGHIFENVI GPKMKLKNKTRILVTHSMSYLPQVDVYIV 840
Db 781 AVYSNADIVLPDPLSADVAHVGHIFENVI GPKMKLKNKTRILVTHSMSYLPQVDVYIV 840
QY 841 MSGGKISEMGSYOBELARDAFAEFLRTYASTOEDDAENGVGSGPKKAKOMENGM 900
Db 841 MSGGKISEMGSYOBELARDAFAEFLRTYASTOEDDAENGVGSGPKKAKOMENGM 900
QY 901 LVYDSAGKQOROLSSSSSYSGDISRHHNSTAELOKAEAKKETWKLMEADKXQOTGVKL 960
Db 901 LVYDSAGKQOROLSSSSSYSGDISRHHNSTAELOKAEAKKETWKLMEADKXQOTGVKL 960
QY 961 SVYWDYMKALGLFISPLSIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTYVRLSVYG 1020
Db 961 SVYWDYMKALGLFISPLSIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTYVRLSVYG 1020
QY 1021 ALGISOGIAVFGSMVASTIGIILASRCLHVDLHSLIRSPMSFFERPSPSNLVNRSKEL 1080
Db 1021 ALGISOGIAVFGSMVASTIGIILASRCLHVDLHSLIRSPMSFFERPSPSNLVNRSKEL 1080
QY 1081 DTVDSMIPEYIKMFMSLFNVIGACIYIILATPIAIIIPPLGLIYEFVORFYVASSROL 1140
Db 1081 DTVDSMIPEYIKMFMSLFNVIGACIYIILATPIAIIIPPLGLIYEFVORFYVASSROL 1140
QY 1141 KRLESVSRSPVYSHFNETLIGSVIRAFEOERPIHQSDIKVDENOKAYVPSIVANRWLA 1200
Db 1141 KRLESVSRSPVYSHFNETLIGSVIRAFEOERPIHQSDIKVDENOKAYVPSIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVIRSRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVIRSRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVEFRNYCLARYEDDLFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVEFRNYCLARYEDDLFVLRHINVTINGG 1320

QY 1321 EKVGIVRTGAGKSLTGLFRINESAGEIIGIDINIAKIGLHDLREKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSLTGLFRINESAGEIIGIDINIAKIGLHDLREKTIIPQDPVLF 1380
QY 1381 SSGIRNMULDPPSQSDEEWTSLSLAHKDPVSALPDLDBHCEAGGNNLSVGQRLVCL 1440
DB 1381 SSGIRNMULDPPSQSDEEWTSLSLAHKDPVSALPDLDBHCEAGGNNLSVGQRLVCL 1440
QY 1441 ARALLRTKILVLEBATAVADLETDDLIQSTIRTOFEDCTVLTJAHRLNTIMDYTRVVL 1500
DB 1441 ARALLRTKILVLEBATAVADLETDDLIQSTIRTOFEDCTVLTJAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIOEYGAESDILLQORGLFYMAKQAGLV 1531
DB 1501 DKGEIOEYGAESDILLQORGLFYMAKQAGLV 1531

RESULT 7
US-08-141-893-2
; Sequence 2, Application US/08141893
; Patent No. 5489519
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P. C.
; APPLICANT: Deele, Roger G.
; TITLE OF INVENTION: MOLITIDRUG RESISTANCE PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/141,893
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923; 08/029,340
; FILING DATE: 27-OCT-1992; 8-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Decont, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: P01-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5149
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-141-893-2

Query Match 79.3%; Score 7849; DB 1; Length 1531,
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNVNTWNTSNPDTKCFONTVLVWPCFYLMACFPFYFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNVNTWNTSNPDTKCFONTVLVWPCFYLMACFPFYFLYLSRH 60
QY 61 DRGIQMTPLNKKTKTALGFLLMTVCMADLFYSFWRSGIFLAPVPLVSPITLIGITTLA 120
DB 61 DRGIQMTPLNKKTKTALGFLLMTVCMADLFYSFWRSGIFLAPVPLVSPITLIGITTLA 120

QY 121 TFLIQLERRKGVSSGIMLTFWLVALVCALAILRSKMTALKEDAQVDLPDITFYFYS 180
DB 121 TFLIQLERRKGVSSGIMLTFWLVALVCALAILRSKMTALKEDAQVDLPDITFYFYS 180
QY 181 LLLIQVLSCPSDPSLFFSETIHDNCPSPSSASFSLRTFFWMTTGLIVGGRPLGSD 240
DB 181 LLLIQVLSCPSDPSLFFSETIHDNCPSPSSASFSLRTFFWMTTGLIVGGRPLGSD 240
QY 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKQPVVYVSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKQPVVYVSSKDPAPKSSKVDANEVEAL 300
QY 301 IVKSPQKEMNPSELFKYITGPGYFLMSPEFKAIHDMLESFPOLKLLIKFVNDYAPD 360
DB 301 IVKSPQKEMNPSELFKYITGPGYFLMSPEFKAIHDMLESFPOLKLLIKFVNDYAPD 360
QY 361 WQGYFYTVLLFVTAQCTQTLVHQYFHCPSGMRKTAIVGAVVRKALVTNSARKSTV 420
DB 361 WQGYFYTVLLFVTAQCTQTLVHQYFHCPSGMRKTAIVGAVVRKALVTNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMWSAPLQVILALYLLMLNGPSVLGAVAVMYLMPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMWSAPLQVILALYLLMLNGPSVLGAVAVMYLMPVN 480
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMSEILNGIKVLKYAMELAFKDYALAIROELKVLK 540
DB 481 AVMAKTKTYQVAHMKSKDNRIKLMSEILNGIKVLKYAMELAFKDYALAIROELKVLK 540
QY 541 KSAVLSVGTFTWCTPFLVALCTFAVYVITDENNIIDAOQAFVSLAFNLIRPLNLPL 600
DB 541 KSAVLSVGTFTWCTPFLVALCTFAVYVITDENNIIDAOQAFVSLAFNLIRPLNLPL 600
QY 601 MVISIIQVASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVENATFTMARSPPT 660
DB 601 MVISIIQVASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVENATFTMARSPPT 660
QY 661 LMGITFSIPEGALVAVVQVCGKSLLSALLAEMDKEGHVAKGSVAVVPOQAMIQND 720
DB 661 LMGITFSIPEGALVAVVQVCGKSLLSALLAEMDKEGHVAKGSVAVVPOQAMIQND 720
QY 721 SLRENILFGCOLLEPPYRSVIAQCALLPDLEILPSGRTIEGEGVNLISGQKQKRVSLAR 780
DB 721 SLRENILFGCOLLEPPYRSVIAQCALLPDLEILPSGRTIEGEGVNLISGQKQKRVSLAR 780
QY 781 AVYENADIYLPDDPLSAVDAVGHIFENYVIGPGMKLNKRIIVTSMSTYLPQVDVIYV 840
DB 781 AVYENADIYLPDDPLSAVDAVGHIFENYVIGPGMKLNKRIIVTSMSTYLPQVDVIYV 840
QY 841 MSGGKISEMGSYQELIARDGAFAEFLRTYASTEQEODAEENGVTGVSQPKAEAKOMENG 900
DB 841 MSGGKISEMGSYQELIARDGAFAEFLRTYASTEQEODAEENGVTGVSQPKAEAKOMENG 900
QY 901 LVYDSAGKOLOROLSSSSSYSGDISRHNSYAELOKBAKKEETWKLMEADKATQGYKL 960
DB 901 LVYDSAGKOLOROLSSSSSYSGDISRHNSYAELOKBAKKEETWKLMEADKATQGYKL 960
QY 961 SVYWDYKKAIGLFTISIFLFMCNHYSAASNYWLSMTDPIVNGTQETHKTVRLSYG 1020
DB 961 SVYWDYKKAIGLFTISIFLFMCNHYSAASNYWLSMTDPIVNGTQETHKTVRLSYG 1020
QY 1021 ALGISOGIAVGYSAVSIIGIILASRCLHVDLHLSILRSPSPFERPDSGNLVNRFSEL 1080
DB 1021 ALGISOGIAVGYSAVSIIGIILASRCLHVDLHLSILRSPSPFERPDSGNLVNRFSEL 1080
QY 1081 DTVDSMIPEVIKMFMSLFNVIGACIVILATPIAIIIPPLGLIYFVQRPVYASRQL 1140
DB 1081 DTVDSMIPEVIKMFMSLFNVIGACIVILATPIAIIIPPLGLIYFVQRPVYASRQL 1140
QY 1141 KRLSVSRSPVYSHPNFTLGVSVIRAPFEEQERTIHOSDLKVDENOKRAYYSIYANRMLA 1200
DB 1141 KRLSVSRSPVYSHPNFTLGVSVIRAPFEEQERTIHOSDLKVDENOKRAYYSIYANRMLA 1200
QY 1201 VRECEVGNCLVFAALPAVISRHSLSAGLVGLSVSYSLQVTTYIINMLVRRMSSEMETNIVA 1260

Db 1201 VRLCEVNCIVLFFALFAVLSRHSLSAGLVLSYSLSQVTTYINMLVRSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPMOIQEAPRPSWPQVREFRNCAVRRDLDFVLRPHINVTNG 1320
Db 1261 VERLKEYSETEKEAPMOIQEAPRPSWPQVREFRNCAVRRDLDFVLRPHINVTNG 1320
Qy 1321 EKVGIVRTGAGKSSLTLGLFRINESAGEIIGININAKIGLHDLFPKTIIPQDPVLF 1380
Db 1321 EKVGIVRTGAGKSSLTLGLFRINESAGEIIGININAKIGLHDLFPKTIIPQDPVLF 1380
Qy 1381 SGIIRKMLDPPSQYSDEEWTSLLELAHKDFVSLPDKLDECAEGGENLSVGQRLVCL 1440
Db 1381 SGIIRKMLDPPSQYSDEEWTSLLELAHKDFVSLPDKLDECAEGGENLSVGQRLVCL 1440
Qy 1441 ARALLRTKIVLDEATPAVDLETDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRYVL 1500
Db 1441 ARALLRTKIVLDEATPAVDLETDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRYVL 1500
Qy 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 8

US-08-463-092B-2
Sequence 2, Application US/08463092B
Patent No. 576680
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-092B-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLDMANTYNTSNDPFKCFONTLVWVPCFYLMACEPFYLYSRH 60
Db 1 MALRGFCSADGSDPLDMANTYNTSNDPFKCFONTLVWVPCFYLMACEPFYLYSRH 60
Qy 61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAVFLVSPTLGITTLLA 120
Db 61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAVFLVSPTLGITTLLA 120
Qy 121 TFLIQLERRKGVSSGIMLTFWVYALCALAIRSKIMTLKEDAOVDLFRDITFFYFYS 180
Db 121 TFLIQLERRKGVSSGIMLTFWVYALCALAIRSKIMTLKEDAOVDLFRDITFFYFYS 180
Qy 181 LLLIQVLSCFSDRSPLEFSETIHDPNPCPSSASFSLRITFMWITGLIVRGYRPLEGSD 240
Db 181 LLLIQVLSCFSDRSPLEFSETIHDPNPCPSSASFSLRITFMWITGLIVRGYRPLEGSD 240
Qy 241 LMSLNKEDTSEQVVPVLYKMKKECAKTRQPKVYVSSKDPAPCKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLYKMKKECAKTRQPKVYVSSKDPAPCKESSKYDANEVEAL 300
Qy 301 IVKSPQKEMNPSLFKVLKTFPGPYFLMSFFPKAIDHLMFSGPOLIKLIKFNVDYAPD 360
Db 301 IVKSPQKEMNPSLFKVLKTFPGPYFLMSFFPKAIDHLMFSGPOLIKLIKFNVDYAPD 360
Qy 361 MGGFYTVLLFPTACLOTVLHOFHICFSGMRKTAIVGAYRKALVTNSRKSTV 420
Db 361 MGGFYTVLLFPTACLOTVLHOFHICFSGMRKTAIVGAYRKALVTNSRKSTV 420
Qy 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILLYLMLNGLPSVLAGVAVMLVAVN 480
Db 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILLYLMLNGLPSVLAGVAVMLVAVN 480
Qy 481 AVNPKTKTYQVAMKSKNRILKAMEINGIVLYKYAMELAFKDYLAIRQELKVLK 540
Db 481 AVNPKTKTYQVAMKSKNRILKAMEINGIVLYKYAMELAFKDYLAIRQELKVLK 540
Qy 541 KSAVLSAVGTFWVCTPEVALCTPAVYVITIDENNILDAQTAFVSLAFNILLRPPNILLP 600
Db 541 KSAVLSAVGTFWVCTPEVALCTPAVYVITIDENNILDAQTAFVSLAFNILLRPPNILLP 600
Qy 601 MVISIIVQASVSLKRLIFLSHELEBDSIERRPVDDGGTNSITVYNAFTWARSDDPT 660
Db 601 MVISIIVQASVSLKRLIFLSHELEBDSIERRPVDDGGTNSITVYNAFTWARSDDPT 660
Qy 661 LMGITTSIPREGALVAVVGVCCKSSLLSALLAEMDKVEGHVAIKSVAVVPOQAWIOND 720
Db 661 LMGITTSIPREGALVAVVGVCCKSSLLSALLAEMDKVEGHVAIKSVAVVPOQAWIOND 720
Qy 721 SLRENTLFGQULEEPPYRSYIOCALLPDEILIPSGGRTTIGKGVNLSCGQORVSLAR 780
Db 721 SLRENTLFGQULEEPPYRSYIOCALLPDEILIPSGGRTTIGKGVNLSCGQORVSLAR 780
Qy 781 AVYSNADIVLFDPLSLAVDAHVGHKIFENYIGKGMKNKTRILLVTHSMYSLPQVDYIIV 840
Db 781 AVYSNADIVLFDPLSLAVDAHVGHKIFENYIGKGMKNKTRILLVTHSMYSLPQVDYIIV 840
Qy 841 MSGGKISEMGSYQELLARDAFAEPLRTYASTQEBDAENGYTYGSGPGEAKQOMENGM 900
Db 841 MSGGKISEMGSYQELLARDAFAEPLRTYASTQEBDAENGYTYGSGPGEAKQOMENGM 900
Qy 901 LVTDGAKQQLQRLSSSSSVSGDISRHHSNTAEIQLAKAEKEFTWKLMEADKXQQTGVKL 960

|||||
Db 901 LVTDSACKOJLOROLSSSSSSSGDLSRHNSIAELQKAEKKEETWKLMEADKAOYQVCL 960
Qy 961 SVYDYMKAJGLFTSFISIFLFCMCHVASLASNYWLSMTDDPIYNGOETHKRLSYG 1020
Db 961 SVYDYMKAJGLFTSFISIFLFCMCHVASLASNYWLSMTDDPIYNGOETHKRLSYG 1020
Qy 1021 ALGISOGIAVGYSMVAISIGGILASRCILAVDLHLHSILSPMSFFERTPSGNLVNRFSEL 1080
Db 1021 ALGISOGIAVGYSMVAISIGGILASRCILAVDLHLHSILSPMSFFERTPSGNLVNRFSEL 1080
Qy 1081 DTUDSMIPBVKMWSLFWNIGACIVLALTPIAIIIPPLGLIYFVQRFYVASSQOL 1140
Db 1081 DTUDSMIPBVKMWSLFWNIGACIVLALTPIAIIIPPLGLIYFVQRFYVASSQOL 1140
Qy 1141 KRLBSVSRSPYSHFNFTLGVSVYIRAFEBQERFHOSDLKVDEKQAYYSIVANRWLA 1200
Db 1141 KRLBSVSRSPYSHFNFTLGVSVYIRAFEBQERFHOSDLKVDEKQAYYSIVANRWLA 1200
Qy 1201 VRLBVCNGCIVLFAALFAVISRHSLSAGLVGLSVSYLQVTTYLNMVLVRMSSEMETNIVA 1260
Db 1201 VRLBVCNGCIVLFAALFAVISRHSLSAGLVGLSVSYLQVTTYLNMVLVRMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKAPWQIOETAPSSWPQVGRVFRNTCLAYREDLDVLRHINVTINGG 1320
Db 1261 VERLKEVSETEKAPWQIOETAPSSWPQVGRVFRNTCLAYREDLDVLRHINVTINGG 1320
Qy 1321 EKVGVIGRTGAGKSLTLGLFRINSAGEIIIDGINIAKIGLHDLRFKITIIPDDPLVF 1380
Db 1321 EKVGVIGRTGAGKSLTLGLFRINSAGEIIIDGINIAKIGLHDLRFKITIIPDDPLVF 1380
Qy 1381 SGLSRMNLDPFSSOYSDSEEWMTSLBLAHUKDFVSALPDKLDHCAEGENLSVGORQLVCL 1440
Db 1381 SGLSRMNLDPFSSOYSDSEEWMTSLBLAHUKDFVSALPDKLDHCAEGENLSVGORQLVCL 1440
Qy 1441 ARALIRKTKIIVLDBATVAVLETTDILQSTIRTOFECCTVLTIAHRIANTIMOTRIVVL 1500
Db 1441 ARALIRKTKIIVLDBATVAVLETTDILQSTIRTOFECCTVLTIAHRIANTIMOTRIVVL 1500
Qy 1501 DKGEIOEYGAPSDLLQORGLFYSMKADGVL 1531
Db 1501 DKGEIOEYGAPSDLLQORGLFYSMKADGVL 1531

RESULT 9

US-08-462-109A-2
Sequence 2, Application US/08462109A
Patent No. 5882875
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08462,109A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992

APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFSADGSDPLMDMNTWNTSNPDFTKCFONTVLVWPCEYLMACPEFFLYLSRH 60
Db 1 MALRGFSADGSDPLMDMNTWNTSNPDFTKCFONTVLVWPCEYLMACPEFFLYLSRH 60
Qy 61 DRGIQNTPLNKTALGFLIMLYCNDLPFSFEBRSRGIFLAVPLVSPLLGITTLLA 120
Db 61 DRGIQNTPLNKTALGFLIMLYCNDLPFSFEBRSRGIFLAVPLVSPLLGITTLLA 120
Qy 121 TFLQLERKRGVSSGIMLTFMLVALCALAIISSKIMTALKEDAOYDLDFDITFYVYS 180
Db 121 TFLQLERKRGVSSGIMLTFMLVALCALAIISSKIMTALKEDAOYDLDFDITFYVYS 180
Qy 181 LLLIQVLSCFSDRSPLESETIHDNPNCPRESSASFLSRITFMWITGLIVRGYROPLEGSD 240
Db 181 LLLIQVLSCFSDRSPLESETIHDNPNCPRESSASFLSRITFMWITGLIVRGYROPLEGSD 240
Qy 241 LMSLNKEDTSQVVPVLYKMKKECATRKQPVVYVSSQDPAPKSSKVDANEVEVAL 300
Db 241 LMSLNKEDTSQVVPVLYKMKKECATRKQPVVYVSSQDPAPKSSKVDANEVEVAL 300
Qy 301 IVKSPQKEMNPSLKVLYKTPGFYELMSFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLKVLYKTPGFYELMSFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360
Qy 361 WQGYFYVLLFVTACLOTVLHQYFHI CFVSGMRIKTAVIGAVVRKALVITNSARKSSTV 420
Db 361 WQGYFYVLLFVTACLOTVLHQYFHI CFVSGMRIKTAVIGAVVRKALVITNSARKSSTV 420
Qy 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOYIALLYLMLNLSGSVLAGVAVMTLMPVN 480
Db 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOYIALLYLMLNLSGSVLAGVAVMTLMPVN 480
Qy 481 AVMAKTKTYOVAAHMKSDNR IKLMEIINIGI KYLKYAMELAFKDKVLAIROBELKYLK 540
Db 481 AVMAKTKTYOVAAHMKSDNR IKLMEIINIGI KYLKYAMELAFKDKVLAIROBELKYLK 540
Qy 541 KSAVLSAVGFTWCTPFLVALCTFAVAVTIDENNIIIDAQTAFFSLALFNILRPLNLTLP 600
Db 541 KSAVLSAVGFTWCTPFLVALCTFAVAVTIDENNIIIDAQTAFFSLALFNILRPLNLTLP 600
Qy 601 MVISIVQASVSLKRLRIFLSHELEPDSIRRPVKGCGTNSITVRNATFTWARSDDPT 660
Db 601 MVISIVQASVSLKRLRIFLSHELEPDSIRRPVKGCGTNSITVRNATFTWARSDDPT 660
Qy 661 LMGITFSPBGAIVAVVQVCGGSSLSLALAMDKVGEHVALIKGSAVYVPOQAMTND 720
Db 661 LMGITFSPBGAIVAVVQVCGGSSLSLALAMDKVGEHVALIKGSAVYVPOQAMTND 720

QY 721 SLRENILFGCOLLEPPYRSVIAOCALLPDLLEILPSGRTEIGEGVNLGGOKORVELAR 780
DB 721 SLRENILFGCOLLEPPYRSVIAOCALLPDLLEILPSGRTEIGEGVNLGGOKORVELAR 780
QY 781 AVYSNADIYLPDDPLSAVDHVGKHEENVIGPKMLKNKTRILVTHSMSTYLPQVDYIIV 840
DB 781 AVYSNADIYLPDDPLSAVDHVGKHEENVIGPKMLKNKTRILVTHSMSTYLPQVDYIIV 840
QY 841 MSGGKISEMGSYOBLARDGAFEEILTYASTEOBDAEENGTVGSGPGKEAQMONGM 900
DB 841 MSGGKISEMGSYOBLARDGAFEEILTYASTEOBDAEENGTVGSGPGKEAQMONGM 900
QY 901 LVYDSAGKOLQORLSSSSSSSGDISRRHNSTAELOKAEKEETWKMMEADKATQGVKL 960
DB 901 LVYDSAGKOLQORLSSSSSSSGDISRRHNSTAELOKAEKEETWKMMEADKATQGVKL 960
QY 961 SVTWDMKAIIGLIFSLIFELFMCNHNVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVTWDMKAIIGLIFSLIFELFMCNHNVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVPGVMAVSGIIGIILASRCLHVDLHSLIRSPMSFEERTPSGMLVNRFSKEL 1080
DB 1021 ALGISOGIAVPGVMAVSGIIGIILASRCLHVDLHSLIRSPMSFEERTPSGMLVNRFSKEL 1080
QY 1081 DTVDMSIPEVIKPMFMSLFNVIACIYILATPIAIIIPPLGLIYFFVORFYVASSROL 1140
DB 1081 DTVDMSIPEVIKPMFMSLFNVIACIYILATPIAIIIPPLGLIYFFVORFYVASSROL 1140
QY 1141 KRLSEVSRSVPYSHFNFTLIGSVYIRA FEEDERFHOSDKVDENOKAYYPSIVANRMLA 1200
DB 1141 KRLSEVSRSVPYSHFNFTLIGSVYIRA FEEDERFHOSDKVDENOKAYYPSIVANRMLA 1200
QY 1201 VRLCEVCNCIVLPALPAVVISRHSLSGLVGLSVSYLOVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVCNCIVLPALPAVVISRHSLSGLVGLSVSYLOVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFRNYCLRYREDLFPVLRHINVTNGG 1320
DB 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFRNYCLRYREDLFPVLRHINVTNGG 1320
QY 1321 EKVGIWRTGAGKSSLTGLFRINESAGEIIGIINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIWRTGAGKSSLTGLFRINESAGEIIGIINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SSGLRNMLDPPSQYSDEEVTSLLEHLKPVSAALPKLDHECAGEGENTSVGQRQVCL 1440
DB 1381 SSGLRNMLDPPSQYSDEEVTSLLEHLKPVSAALPKLDHECAGEGENTSVGQRQVCL 1440
QY 1441 ARALLRKTILVDEATAVADLETDDLIQSTIRPOFEDCTVLTAAHRLNTIMOTRYIVL 1500
DB 1441 ARALLRKTILVDEATAVADLETDDLIQSTIRPOFEDCTVLTAAHRLNTIMOTRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 10
US-08-460-907B-2
Sequence 2, Application US/08460907B
Patent No. 5891724
GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
RESISTANCE ON A CELL.
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA

ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-2
Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALRGFCSDGSDPLMDMNTVNTSNPDTKCFQNTVLVWPQFYLMACPFYFLYSRH 60
DB 1 MALRGFCSDGSDPLMDMNTVNTSNPDTKCFQNTVLVWPQFYLMACPFYFLYSRH 60
QY 61 DRGIQWTPLNKTKTALGFLMIVCWADLFYSWERSRGIFLAPVPLVSTLLGITTLA 120
DB 61 DRGIQWTPLNKTKTALGFLMIVCWADLFYSWERSRGIFLAPVPLVSTLLGITTLA 120
QY 121 TFLIOLERRKGVSSGMLTFMIVCALIIRSKIMTALKBDAQVDLFRDITFYVYS 180
DB 121 TFLIOLERRKGVSSGMLTFMIVCALIIRSKIMTALKBDAQVDLFRDITFYVYS 180
QY 181 LLLIQLVLSGFSRSPFLSFETHDPPNCPSSASFLSRITFMWITGLIVRGYQPLEGSD 240
DB 181 LLLIQLVLSGFSRSPFLSFETHDPPNCPSSASFLSRITFMWITGLIVRGYQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKYVYSSGDPQOPKSSRYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKYVYSSGDPQOPKSSRYDANEVEAL 300
QY 301 IVKSPQKEMNPSEFKVLVYKTFPGPYFLMSFFKAIHDLMPGSGQILKLKFNDRKAPD 360
DB 301 IVKSPQKEMNPSEFKVLVYKTFPGPYFLMSFFKAIHDLMPGSGQILKLKFNDRKAPD 360
QY 361 WGGYFTVLLFVATCIQTLVLAHQYFHCYVSGRIRITAVIGAVYRVALVITNSAKSSTV 420
DB 361 WGGYFTVLLFVATCIQTLVLAHQYFHCYVSGRIRITAVIGAVYRVALVITNSAKSSTV 420

QY 421 GEIYNLMSVDAQRFMDLATTYINMISAPLOVITLALYLLMLNGSVIAGVAMVLMVAVN 480
DB 421 GEIYNLMSVDAQRFMDLATTYINMISAPLOVITLALYLLMLNGSVIAGVAMVLMVAVN 480
QY 481 AVAMAKTKTYOVAMKSKDNRIKIMNELINGIKYLKLYAMELAKDKVLAIROBELVYIK 540
DB 481 AVAMAKTKTYOVAMKSKDNRIKIMNELINGIKYLKLYAMELAKDKVLAIROBELVYIK 540
QY 541 KSAVLSAVGTFTWCTEPLVALCTFAVYVITDENNILLDAQTAFAVSLAFNLRPELNLIP 600
DB 541 KSAVLSAVGTFTWCTEPLVALCTFAVYVITDENNILLDAQTAFAVSLAFNLRPELNLIP 600
QY 601 MVTSSIVQASVSLKRLRIFLSHELEBPDSIRRPVKOGGTSITVRNATFTWARSDDPT 660
DB 601 MVTSSIVQASVSLKRLRIFLSHELEBPDSIRRPVKOGGTSITVRNATFTWARSDDPT 660
QY 661 LINGITFSPGALVAVVGVGCGKSLSLALLAMDKVEGVALKGSVAVVPOQAMQND 720
DB 661 LINGITFSPGALVAVVGVGCGKSLSLALLAMDKVEGVALKGSVAVVPOQAMQND 720
QY 721 SLRENIIFGCOLLEPYRYSVIOACALLPDLRLPSGDRTEIGEGVNLSSGQKORVSLAR 780
DB 721 SLRENIIFGCOLLEPYRYSVIOACALLPDLRLPSGDRTEIGEGVNLSSGQKORVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDAHVGKHFENVIQPKGMLKNKTRIIIVTHSMYSILPOVDVIIV 840
DB 781 AVYSNADIYLFDDPLSAVDAHVGKHFENVIQPKGMLKNKTRIIIVTHSMYSILPOVDVIIV 840
QY 841 MSGGKISBMSGYOBLLARDGAFARLRTYASTEOQDAEENGVTGVSGPEAKOMENGM 900
DB 841 MSGGKISBMSGYOBLLARDGAFARLRTYASTEOQDAEENGVTGVSGPEAKOMENGM 900
QY 901 LVTDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKKEETKLMLEADKOTGVYKL 960
DB 901 LVTDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKKEETKLMLEADKOTGVYKL 960
QY 961 SVYWDYMAKIGLFSISLIFLMCNHVSALASNTWLSMTDPIVNGTQEHTKVRLSYVG 1020
DB 961 SVYWDYMAKIGLFSISLIFLMCNHVSALASNTWLSMTDPIVNGTQEHTKVRLSYVG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHNDLHLSILASPMSPFERTSGULNRFSEL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHNDLHLSILASPMSPFERTSGULNRFSEL 1080
QY 1081 DTVYSMIPEVIMKMGSLFNVIYACIVILLATPIAIIIPPLGIYFFQRFYVYASSROL 1140
DB 1081 DTVYSMIPEVIMKMGSLFNVIYACIVILLATPIAIIIPPLGIYFFQRFYVYASSROL 1140
QY 1141 KRLSVSHSPVYSHFNETLLGVSIVIRAFEEQERFIHQSDLKVDENOKAYPYSIVANRWLA 1200
DB 1141 KRLSVSHSPVYSHFNETLLGVSIVIRAFEEQERFIHQSDLKVDENOKAYPYSIVANRWLA 1200
QY 1201 VRLFCVNGCIVLPALFAVIRSHSISAGLVGSYSIQVTTYLNLVYRMSSEMETNIVA 1260
DB 1201 VRLFCVNGCIVLPALFAVIRSHSISAGLVGSYSIQVTTYLNLVYRMSSEMETNIVA 1260
QY 1261 VERLEKESYETEKEAPWMOIETRPPSPWQGRVREFRANCXKREDDVDVLRHINVTINGG 1320
DB 1261 VERLEKESYETEKEAPWMOIETRPPSPWQGRVREFRANCXKREDDVDVLRHINVTINGG 1320
QY 1321 EKVGIYGTGAGKSSLTJGLFRINSAGEI11IDGINIAKIGLHDLRFKTI11PODPVLF 1380
DB 1321 EKVGIYGTGAGKSSLTJGLFRINSAGEI11IDGINIAKIGLHDLRFKTI11PODPVLF 1380
QY 1381 SSGSLRMINIDPPSQYSDEEVTSLLELAHKDFVSALPDKLDHCCAGGENTLSSGQOLVCL 1440
DB 1381 SSGSLRMINIDPPSQYSDEEVTSLLELAHKDFVSALPDKLDHCCAGGENTLSSGQOLVCL 1440
QY 1441 ABALLRKTKIIVLDATAANDLETDLLQSTIRTFQEDCTVLTINHRINTIMDYRIVVL 1500
DB 1441 ABALLRKTKIIVLDATAANDLETDLLQSTIRTFQEDCTVLTINHRINTIMDYRIVVL 1500
QY 1501 DKGEIOEYGAPSDLLQGRGLFYSMAKDAGLV 1531

DB 1501 DKGEIOEYGAPSDLLQGRGLFYSMAKDAGLV 1531
RESULT 11
US-08-463-179A-2
Sequence 2, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Decont, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-2
Query Match 79.3%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALNGFSGADSDPLMDNMVNTSNPDTKCFONTVLMWPCFYLMACPFYFLYSRH 60
DB 1 MALNGFSGADSDPLMDNMVNTSNPDTKCFONTVLMWPCFYLMACPFYFLYSRH 60
QY 61 DRGIYQMTPLNKTKTALGFLMLIVCNADLFYSFWERSRGIFLAEVFLVSPTLIGITTLA 120
DB 61 DRGIYQMTPLNKTKTALGFLMLIVCNADLFYSFWERSRGIFLAEVFLVSPTLIGITTLA 120
QY 121 TFLIQLERRKGVSGGIMLTWLVALVCAALILRSKIMTALKEDAQVDLFPDITFYVYFS 180
DB 121 TFLIQLERRKGVSGGIMLTWLVALVCAALILRSKIMTALKEDAQVDLFPDITFYVYFS 180
QY 181 LLLIQLVISCFSDSSPLFSETHDNPCESSASFLSITFWMTTGLVREGYRPLBSSD 240
DB 181 LLLIQLVISCFSDSSPLFSETHDNPCESSASFLSITFWMTTGLVREGYRPLBSSD 240

QY 241 LMSLNKEDTSBOVPLVVKMKKECAKTRKOPVKVYSSKDPAPQPKSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSBOVPLVVKMKKECAKTRKOPVKVYSSKDPAPQPKSSKVDANEVEAL 300
 QY 301 IVKSPQKEMNPISLFKVLKTFPGPYFLMSFPFKAIHDLMPFSGPOLIKLIFVNDTRAPD 360
 DB 301 IVKSPQKEMNPISLFKVLKTFPGPYFLMSFPFKAIHDLMPFSGPOLIKLIFVNDTRAPD 360
 QY 361 WQGFYVVLVFNACLOTLVHQFPHICFVSGMGIKTAIVIGAVYRKALVTINARKSTV 420
 DB 361 WQGFYVVLVFNACLOTLVHQFPHICFVSGMGIKTAIVIGAVYRKALVTINARKSTV 420
 QY 421 GEIYNLMSVDAOREFMDLATYINMWSAPLOVITLMLTNGPSVLAGVAVWVLAMPVN 480
 DB 421 GEIYNLMSVDAOREFMDLATYINMWSAPLOVITLMLTNGPSVLAGVAVWVLAMPVN 480
 QY 481 AVAMKTKTYQVAMKSKDNRIKLMNEILNGIKVLKYLAMELAFKDYLAIRBELKVLK 540
 DB 481 AVAMKTKTYQVAMKSKDNRIKLMNEILNGIKVLKYLAMELAFKDYLAIRBELKVLK 540
 QY 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVITDENNIIDAQTAFLALFNILRPLNLP 600
 DB 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVITDENNIIDAQTAFLALFNILRPLNLP 600
 QY 601 MWISSIVQASVSLKRLIFLSHELEPDSIERRPVKGGSNSTYVNAFTFARSDPT 660
 DB 601 MWISSIVQASVSLKRLIFLSHELEPDSIERRPVKGGSNSTYVNAFTFARSDPT 660
 QY 661 LMGITFESIPGALVAVVGCGCKSLLSALLAEMDVEGHVAKGSVAVYVPOQAWIOND 720
 DB 661 LMGITFESIPGALVAVVGCGCKSLLSALLAEMDVEGHVAKGSVAVYVPOQAWIOND 720
 QY 721 SLRBNILFGCQLEPPYRSVYQACALLPDLILPDSGRTEIGEKNVLSGGOKORVSLAR 780
 DB 721 SLRBNILFGCQLEPPYRSVYQACALLPDLILPDSGRTEIGEKNVLSGGOKORVSLAR 780
 QY 781 AVYSNNDIYLPDPLSAVDHVGKHEENYIGPCKMLKNTRLIYTHSMVTLPOVDYIIV 840
 DB 781 AVYSNNDIYLPDPLSAVDHVGKHEENYIGPCKMLKNTRLIYTHSMVTLPOVDYIIV 840
 QY 841 MSGGKISEMGSYQELARDGAFAEFLRTVASTEOBQDAEENGVTGSGPGKEAKOMENG 900
 DB 841 MSGGKISEMGSYQELARDGAFAEFLRTVASTEOBQDAEENGVTGSGPGKEAKOMENG 900
 QY 901 LVTDSAGKOLQROULSSSSSYSGDISRHNSHTAEIQAQAKKEETWKLMEADKATQGVKL 960
 DB 901 LVTDSAGKOLQROULSSSSSYSGDISRHNSHTAEIQAQAKKEETWKLMEADKATQGVKL 960
 QY 961 SYVWDVWKATGLFISFLSIFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
 DB 961 SYVWDVWKATGLFISFLSIFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
 QY 1021 ALGISOGIAVFGYSMAVISGIIASRCIAYDLHSLIRSPMSPERTPSGNLVNRFSEKL 1080
 DB 1021 ALGISOGIAVFGYSMAVISGIIASRCIAYDLHSLIRSPMSPERTPSGNLVNRFSEKL 1080
 QY 1081 DTVDMSMPEVYIKMFGSLFNVIIGACIYILLATPIAIIIPPLGIYFVORFYVASSROL 1140
 DB 1081 DTVDMSMPEVYIKMFGSLFNVIIGACIYILLATPIAIIIPPLGIYFVORFYVASSROL 1140
 QY 1141 KRLESVRSRPSVSHFNETLLGVSVIRAFEEQERFIHOSDKVDENQAKYPSIYANWLA 1200
 DB 1141 KRLESVRSRPSVSHFNETLLGVSVIRAFEEQERFIHOSDKVDENQAKYPSIYANWLA 1200
 QY 1201 VRLCEVNCIYVLPALFAVISRHSLSAGLVGLSVSYSLQVTTYANWLVRRSSMETIYVA 1260
 DB 1201 VRLCEVNCIYVLPALFAVISRHSLSAGLVGLSVSYSLQVTTYANWLVRRSSMETIYVA 1260
 QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVAFENYCLRYREDLDFVLRIHINVTINGG 1320
 DB 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVAFENYCLRYREDLDFVLRIHINVTINGG 1320
 QY 1321 EKVGIVGTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLAFKLTITIPQDVLV 1380

DB 1321 EKVGIVGTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLAFKLTITIPQDVLV 1380
 QY 1381 SCSLRNMLDPPSOYSDEEVTSLSLAHLKDPVSALPKDLHCEAGGENTSVGQROLVCL 1440
 DB 1381 SCSLRNMLDPPSOYSDEEVTSLSLAHLKDPVSALPKDLHCEAGGENTSVGQROLVCL 1440
 QY 1441 ARALLRKTILVDEAFAADLETFDDLIQSTIRTOPEBDCTVLTARHLNTIMOTRYIVL 1500
 DB 1441 ARALLRKTILVDEAFAADLETFDDLIQSTIRTOPEBDCTVLTARHLNTIMOTRYIVL 1500
 QY 1501 DKGEIOEYCAPSDILOORGLFYSGMAXDAGLV 1531
 DB 1501 DKGEIOEYCAPSDILOORGLFYSGMAXDAGLV 1531
 RESULT 12
 US-08-461-384B-2
 : Sequence 2, Application US/08461384B
 : Patent No. 6025473
 : GENERAL INFORMATION:
 : APPLICANT: Cole, Susan P.C.
 : APPLICANT: Deeley, Roger G.
 : TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
 : NUMBER OF SEQUENCES: 10
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
 : STREET: Queen's University at Kingston
 : CITY: Kingston
 : STATE: Ontario
 : COUNTRY: CANADA
 : ZIP: K7L 3N6
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: ASCII text
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/461,384B
 : FILING DATE: 05-JUN-95
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/966,923
 : FILING DATE: 27-OCT-1992
 : APPLICATION NUMBER: 08/029,340
 : FILING DATE: 8-MAR-1993
 : APPLICATION NUMBER: 08/141,893
 : FILING DATE: 26-OCT-1993
 : APPLICATION NUMBER: 08/407,207
 : FILING DATE: 20-MAR-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Steeg, Carol Miernicki
 : REGISTRATION NUMBER: 39,539
 : REFERENCE/DOCKET NUMBER: 01547
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (613) 545-2342
 : TELEFAX: (613) 545-6853
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1531 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-08-461-384B-2
 Query Match 79.3%; Score 7849; DB 2; Length 1531;
 Best Match Similarity 99.9%; Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MALRGFCSDAGSDPLMDMNTVNTSNPDFTKCFONTLVVWPCFYIMACEPFYFLYSR 60
 DB 1 MALRGFCSDAGSDPLMDMNTVNTSNPDFTKCFONTLVVWPCFYIMACEPFYFLYSR 60
 QY 61 DRGYIGMTPLNKTKTALGLFLMIVCWADLFYFWRERSRGIFLAPVFLVSPFTLLGITTLA 120

```

Db      61 DKGITQMTPLNKTKTALGFLMTIVCMADLFYSFMRSGKIFLAVFLVSPILLGITLLA 120
Qy      121 TPLIQERRRGVSSGIMLTFMVALVCAALIRSKIMTALKEADVDLDFDITFEYYS 180
Db      121 TPLIQERRRGVSSGIMLTFMVALVCAALIRSKIMTALKEADVDLDFDITFEYYS 180
Qy      181 LLLIQVLSCFSDRSPLFSETIHDNPPCPESSASFLSKITFMTITGLVRCGRPLBGS 240
Db      181 LLLIQVLSCFSDRSPLFSETIHDNPPCPESSASFLSKITFMTITGLVRCGRPLBGS 240
Qy      241 LMSLNKEDTSQVVPVLYKWKKECAKTRKOPVYVYSSKDPAPKSSSKVDANEVEAL 300
Db      241 LMSLNKEDTSQVVPVLYKWKKECAKTRKOPVYVYSSKDPAPKSSSKVDANEVEAL 300
Qy      301 IVSPKEMNPSSLKVLKTFEGPYFLMSFFKAIHDLMMFSGPOLKLLIFVNDTAPD 360
Db      301 IVSPKEMNPSSLKVLKTFEGPYFLMSFFKAIHDLMMFSGPOLKLLIFVNDTAPD 360
Qy      361 WQGYFYVLLFVTACLOTLVLHQYFHLCEVSGMRIKTAVIGAVYRKALVITNSARKSTV 420
Db      361 WQGYFYVLLFVTACLOTLVLHQYFHLCEVSGMRIKTAVIGAVYRKALVITNSARKSTV 420
Qy      421 GEIYNLMSVDAQRMDLATTYNNMTWSAPLOYITALLYLMLNLPGRSVLAGVAVMVLMPVN 480
Db      421 GEIYNLMSVDAQRMDLATTYNNMTWSAPLOYITALLYLMLNLPGRSVLAGVAVMVLMPVN 480
Qy      481 AVNMAKTKTYOVAMHKSNDRIKLMNEILNGIKVLKYAMELAFDKYLAIRSEELKYLK 540
Db      481 AVNMAKTKTYOVAMHKSNDRIKLMNEILNGIKVLKYAMELAFDKYLAIRSEELKYLK 540
Qy      541 KSAVLSAVGTFTWCTPFLVALCTFAVYVYIDENNILDAQTAFLVALFNILRPNILP 600
Db      541 KSAVLSAVGTFTWCTPFLVALCTFAVYVYIDENNILDAQTAFLVALFNILRPNILP 600
Qy      601 MVTISIVQASVSLKRLRIFLSHEBIEPDSIERRPYKOGGINSITVRNATFTWASDPT 660
Db      601 MVTISIVQASVSLKRLRIFLSHEBIEPDSIERRPYKOGGINSITVRNATFTWASDPT 660
Qy      661 LINGITFSLPEGALVAVVQVCGKSLLSALLAENDKYEGHVAIKGSAAVYPOQAMIOND 720
Db      661 LINGITFSLPEGALVAVVQVCGKSLLSALLAENDKYEGHVAIKGSAAVYPOQAMIOND 720
Qy      721 SLRENILFGCOLBEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQOKRVSLAR 780
Db      721 SLRENILFGCOLBEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQOKRVSLAR 780
Qy      781 AVYSNADIYLFDDPLSAVDHVKHIFENVIQPKMLKMKTRILVTHSMSTLPQVDVITV 840
Db      781 AVYSNADIYLFDDPLSAVDHVKHIFENVIQPKMLKMKTRILVTHSMSTLPQVDVITV 840
Qy      841 MSGGKISGMSYOEILARDGAFABELRTYASTEOBDAENGVMGVSGPGKEAKOMENGM 900
Db      841 MSGGKISGMSYOEILARDGAFABELRTYASTEOBDAENGVMGVSGPGKEAKOMENGM 900
Qy      901 LVYDSAGKOLOROLSSSSSYSGDISRHHNSTAELOKAKAKEETKLMLEADKAOQGYKL 960
Db      901 LVYDSAGKOLOROLSSSSSYSGDISRHHNSTAELOKAKAKEETKLMLEADKAOQGYKL 960
Qy      961 SVYWDYMKAIQLFISFLSIFLPMCNHVSALASNTWLSMTDDPIVNGTQEHTKVRLSYG 1020
Db      961 SVYWDYMKAIQLFISFLSIFLPMCNHVSALASNTWLSMTDDPIVNGTQEHTKVRLSYG 1020
Qy      1021 ALGISOGIAVFGYMAVSIIGIILASRCHVLDLHLSILSPSPFERTSGULVNFSEKL 1080
Db      1021 ALGISOGIAVFGYMAVSIIGIILASRCHVLDLHLSILSPSPFERTSGULVNFSEKL 1080
Qy      1081 DTVDSMTPEVIMKMGSLFNVIYACIVILLATPIAIIIPPLGILYFVQRFYVASSQOL 1140
Db      1081 DTVDSMTPEVIMKMGSLFNVIYACIVILLATPIAIIIPPLGILYFVQRFYVASSQOL 1140
Qy      1141 KRLSVSRSPYVSHNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200

```

```

Db      1141 KRLSVSRSPYVSHNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
Qy      1201 VRLCEVGNICIVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYLNMVLRMSSEMETNIVA 1260
Db      1201 VRLCEVGNICIVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYLNMVLRMSSEMETNIVA 1260
Qy      1261 VERLKEYSETEKEAPWIOETAPSSWPQVGRVEFRNYCLRYRDLDFVLRLHINVTINGG 1320
Db      1261 VERLKEYSETEKEAPWIOETAPSSWPQVGRVEFRNYCLRYRDLDFVLRLHINVTINGG 1320
Qy      1321 EKVGIVERTAGKSSLTGLFRINESAGEIILIDGINIAKIGLHDLAFKTIIPRODPVLF 1380
Db      1321 EKVGIVERTAGKSSLTGLFRINESAGEIILIDGINIAKIGLHDLAFKTIIPRODPVLF 1380
Qy      1381 SGLSRMLDPPSQYSDSEVWTSLELAHKDFVSAIPDKLDHECAGGENLSVQOROLVCL 1440
Db      1381 SGLSRMLDPPSQYSDSEVWTSLELAHKDFVSAIPDKLDHECAGGENLSVQOROLVCL 1440
Qy      1441 ARALLRRTKILVDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db      1441 ARALLRRTKILVDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Qy      1501 DKGIOEYGA PSDLQORGLFYMAKADGLV 1531
Db      1501 DKGIOEYGA PSDLQORGLFYMAKADGLV 1531

```

```

RESULT 13
US-08-407-207A-2
; Sequence 2, Application US/08407207A
; Patent No. 6063621
; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P. C.
; TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,207A
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: 01512
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-407-207A-2

```

Query Match 79.3%; Score 7849; DB 2; Length 1531;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MALRGFCGADSGDPLMDNNVTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFYLRLSRH 60
DB 1 MALRGFCGADSGDPLMDNNVTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFYLRLSRH 60
QY 61 DRGYIOMPLNKTAKALGFLIMIVCMADLFYSFMRSRGIFLAPFLVSPFLIGTTTLA 120
DB 61 DRGYIOMPLNKTAKALGFLIMIVCMADLFYSFMRSRGIFLAPFLVSPFLIGTTTLA 120
QY 121 TFLIOERKRGVSSGIMLTFMLVALVICALIIRSKITATLKEQVDFRDIITYYVES 180
DB 121 TFLIOERKRGVSSGIMLTFMLVALVICALIIRSKITATLKEQVDFRDIITYYVES 180
QY 181 LLLIQLVLSCFSDRSPLFSETHDBNCPESASASFLSRITFWITGLIVRGYRQPLBGSD 240
DB 181 LLLIQLVLSCFSDRSPLFSETHDBNCPESASASFLSRITFWITGLIVRGYRQPLBGSD 240
QY 241 LMSLNKETSQVVPVLVKNWKKCEAKTRKQPVKVVYSSKDPAPKXSSKYDANEVEAL 300
DB 241 LMSLNKETSQVVPVLVKNWKKCEAKTRKQPVKVVYSSKDPAPKXSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFLVKYKTPGYFLMSFFPKAIHDLMMPSGPOILKLIKFNVDTKAD 360
DB 301 IVKSPQKEMNPSLFLVKYKTPGYFLMSFFPKAIHDLMMPSGPOILKLIKFNVDTKAD 360
QY 361 WQGYFYTLFVTAQTLQTLVLAHQYFHLICFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
DB 361 WQGYFYTLFVTAQTLQTLVLAHQYFHLICFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAQRFMDLATYINMISAPLQVILALYILMLNGSVLAGAVWVLAVPN 480
DB 421 GEIYNLMSVDAQRFMDLATYINMISAPLQVILALYILMLNGSVLAGAVWVLAVPN 480
QY 481 AVAMKRTKYOVAMHMSKONRIKLMNEILNGIKYKUYAMELAPKDKTALROEELKYLK 540
DB 481 AVAMKRTKYOVAMHMSKONRIKLMNEILNGIKYKUYAMELAPKDKTALROEELKYLK 540
QY 541 KSAVLASAVGTFTWCTPFLVALCTFAVYVITDENNILDQAQFVSLAFNLIRFPLNLP 600
DB 541 KSAVLASAVGTFTWCTPFLVALCTFAVYVITDENNILDQAQFVSLAFNLIRFPLNLP 600
QY 601 MWISSIYOASVSLKRLRIFLSHBELEPDSIRRPVKDGGTNSITVRNATFTWASDPT 660
DB 601 MWISSIYOASVSLKRLRIFLSHBELEPDSIRRPVKDGGTNSITVRNATFTWASDPT 660
QY 661 LINGITFSPREGALVAVVGCGCKSSLSLALLAEMDKVEGHVAKGSVAAYVPOQAMOND 720
DB 661 LINGITFSPREGALVAVVGCGCKSSLSLALLAEMDKVEGHVAKGSVAAYVPOQAMOND 720
QY 721 SLRENILFGCOLLEPPYRSVIOACALLPDLILPSGRTEIGEKNVLSGQXORVSLAR 780
DB 721 SLRENILFGCOLLEPPYRSVIOACALLPDLILPSGRTEIGEKNVLSGQXORVSLAR 780
QY 781 AVYSNADIYLPDPLSAVDAAHVGHIFENVIQPKMLKNKTRILVTHSMSTLPQVDVILV 840
DB 781 AVYSNADIYLPDPLSAVDAAHVGHIFENVIQPKMLKNKTRILVTHSMSTLPQVDVILV 840
QY 841 MSGGKISMGSYOELLARDDGAPAFILRTYASTBOBOAENGUNGVSGPGEAOMENGM 900
DB 841 MSGGKISMGSYOELLARDDGAPAFILRTYASTBOBOAENGUNGVSGPGEAOMENGM 900
QY 901 LVYDSAGKOLOROLSSSSSYSGDISRHNSTAELQKAEKKEFTWKLMEADKAOTGOVKL 960
DB 901 LVYDSAGKOLOROLSSSSSYSGDISRHNSTAELQKAEKKEFTWKLMEADKAOTGOVKL 960
QY 961 SYVYDYNKALIGLTFISLSTFLPKCNHVSALASNTWLSMTDDPIVNGTOEHTKRLSYVG 1020
DB 961 SYVYDYNKALIGLTFISLSTFLPKCNHVSALASNTWLSMTDDPIVNGTOEHTKRLSYVG 1020

```

```

QY 1021 ALGISOGIAVFGYMAVSIIGILASRCILNDLHSLIIRSPMSFFERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGILASRCILNDLHSLIIRSPMSFFERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMIPEVIMKFMGSLFENVIGACIVILLATPIAIIIPPLGLIYFFVQGFYVASSRQL 1140
DB 1081 DTVDSMIPEVIMKFMGSLFENVIGACIVILLATPIAIIIPPLGLIYFFVQGFYVASSRQL 1140
QY 1141 KRLESVSRSPYSHFNELTGVSVIRAPFEQERTIHOSDLKVDENOKRAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNELTGVSVIRAPFEQERTIHOSDLKVDENOKRAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCLVFPALFAVIRSHLSAGLVGLSVYSIQVTTYINMLVRSSEMETNIVA 1260
DB 1201 VRLCEVGNCLVFPALFAVIRSHLSAGLVGLSVYSIQVTTYINMLVRSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIQTETAPSSWPQVGRVBFRRNYCLYREDLDFVLRIHIVTINGG 1320
DB 1261 VERLKEYSETEKAPMOIQTETAPSSWPQVGRVBFRRNYCLYREDLDFVLRIHIVTINGG 1320
QY 1321 EKVGIVGRTGAGKSLTLGLFRINESABGEIINDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTLGLFRINESABGEIINDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRMLNDPFSQYSDSEVWTSLELAHLKDFVSALPKLDHECABGEGENLSVGOROLVCL 1440
DB 1381 SGSLRMLNDPFSQYSDSEVWTSLELAHLKDFVSALPKLDHECABGEGENLSVGOROLVCL 1440
QY 1441 ABALLRKTILVDEATAVADLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ABALLRKTILVDEATAVADLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYMAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDLLQORGLFYMAKDAGLV 1531

```

```

RESULT 14
US-08-463-092B-6
; Sequence 6, Application US/08463092B
; Patent No. 5766880
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P. C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,092B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893

```

FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Mternicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-092B-6

Query Match 70.7%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTVLWVPCFYLMACPFYFLYLSRH 60
DB 1 MALRSCFSDGSDPLMDNVTWNTSNPDFTKCFQNTVLWVPCFYLMACPFYFLYLSRH 60
QY 61 DRGYQMTPLNKTALGFLMIVCMADLFYSFWERSGIGLAPFLVSPITLGLTTTLA 120
DB 61 DRGYQMTPLNKTALGFLMIVCMADLFYSFWERSGIGLAPFLVSPITLGLTTTLA 120
QY 121 TFLQLBRKRVQSSGIMLTFMLVALVCLALIRSKITWALKEDAVDLFRDITFYVFS 180
DB 121 TFLQLBRKRVQSSGIMLTFMLVALVCLALIRSKITWALKEDAVDLFRDITFYVFS 180
QY 181 LLLIQLVLSCFPSDRSPLFSETIHDNPPCPESSASFLSRIITFWITGLIVRGYRPLEGSD 240
DB 181 LVLVOLVLSCFPSDRSPLFSETIHDNPPCPESSASFLSRIITFWITGLIVRGYRPLEGSD 240
QY 241 LMSLNETSEQVNVPLVKNKKKCAKTRKQPVKVVYS- SKDPAOPKSSKVDANEVEA 299
DB 241 LMSLNETSEQVNVPLVKNKKKCAKTRKQPVKVVYS- SKDPAOPKSSKVDANEVEA 299
QY 300 LIVSPKEMNPSLEKVLVYKTFGPFELMSFPFKALHDLMMFSGPOILKLLKFNVDYAP 359
DB 300 LIVSPKEMNPSLEKVLVYKTFGPFELMSFPFKALHDLMMFSGPOILKLLKFNVDYAP 359
QY 360 DMQGYFTVLLFVTAQTLVLHQYFHIQVSGMRKTAIVIGAVYRKALVITNSARKSST 419
DB 360 DMQGYFTVLLFVTAQTLVLHQYFHIQVSGMRKTAIVIGAVYRKALVITNSARKSST 419
QY 420 VGEIYNLMSVDAQRMDLATTYINMTWSAPLOYITALYLLMLNGSVTLAGVAVMLMPL 479
DB 420 VGEIYNLMSVDAQRMDLATTYINMTWSAPLOYITALYLLMLNGSVTLAGVAVMLMPL 479
QY 480 NAVNMMKTKTYOVAMKSKDNRIKLMNELLNGIKYLKLYAMELAKDKVLAIRSEELVYL 539
DB 480 NAVNMMKTKTYOVAMKSKDNRIKLMNELLNGIKYLKLYAMELAKDKVLAIRSEELVYL 539
QY 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVTIDENNIIIDAQTAFAVSLAFNIRPLNIL 599
DB 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVTIDENNIIIDAQTAFAVSLAFNIRPLNIL 599
QY 600 PMVTSISIVQASVSLKRLRIFLSHEELEPDSIRRPVQGGGNSITVNNATFTMARSDP 659
DB 600 PMVTSISIVQASVSLKRLRIFLSHEELEPDSIRRPVQGGGNSITVNNATFTMARSDP 659
QY 660 TLNGITFSIPRGALVAVVGVQCGKSSLLSLALAMDVEGHVALIKGSAAVYPOQAMION 719
DB 660 TLNGITFSIPRGALVAVVGVQCGKSSLLSLALAMDVEGHVALIKGSAAVYPOQAMION 719

QY 720 DSLRENTLFGCCLEPPYRYVIAQCALLPDLEILLPSGDRTIEIGKGVNLSGGQQRVSLA 779
DB 720 DSLRENTLFGHPLOENNYKAMKMACALLPDLEILLPSGDRTIEIGKGVNLSGGQQRVSLA 779
QY 780 RAVYSNADIVLFDPLSAVDAAHVGKHIPENVIGPKGMLKXNTRILVTHSMGYLPOVDYI 839
DB 780 RAVYSNADIVLFDPLSAVDAAHVGKHIPENVIGPKGMLKXNTRILVTHSMGYLPOVDYI 839
QY 840 VMSGKISEMGSYOELLARDGAPAEFLRTYASTEODAEENGTVGVSPPKEAKOMENG 899
DB 840 VMSGKISEMGSYOELLARDGAPAEFLRTYASTEODAEENGTVGVSPPKEAKOMENG 899
QY 900 MLVTDASGKQIQRQLSSSSSYSGDI SRHNS TALQVLEAKKETWKLMEADKATQGVK 959
DB 900 MLVTDASGKQIQRQLSSSSSYSGDI SRHNS TALQVLEAKKETWKLMEADKATQGVK 959
QY 960 LSVYMDVWKAIQLFISFISIFLPMCNHVSALASNYMLSLMTDD-PIVNGTEHKVRLSV 1018
DB 960 LSVYMDVWKAIQLFISFISIFLPMCNHVSALASNYMLSLMTDDPIVNGTEHKVRLSV 1018
QY 1019 YGALGISQGLAVFGYSMAVSIQGLIASRCLHVDLHLSILSPMSFERTPSGNLVNRPSK 1078
DB 1019 YGALGISQGLAVFGYSMAVSIQGLIASRCLHVDLHLSILSPMSFERTPSGNLVNRPSK 1078
QY 1079 ELDTVDKMIPEVYKMFNGSLFNVIGACTVILATPIAIIIPVGLIYFYQRYVASSR 1138
DB 1079 ELDTVDKMIPEVYKMFNGSLFNVIGACTVILATPIAIIIPVGLIYFYQRYVASSR 1138
QY 1139 QLRKLESYSRSPVYSHNETLIGSVIRAPEEORFIHQSDLKVDENOKAYPSTIVANRW 1198
DB 1139 QLRKLESYSRSPVYSHNETLIGSVIRAPEEORFIHQSDLKVDENOKAYPSTIVANRW 1198
QY 1199 LAVRLKCYGNCIVLPAALFAVISRHSLSAGLVGHSVSLQVTTYLNMVMSSEMETNI 1258
DB 1199 LAVRLKCYGNCIVLPAALFAVISRHSLSAGLVGHSVSLQVTTYLNMVMSSEMETNI 1258
QY 1259 VAVRLKESYSTEKBAWQIOETAPPSGWPQVGYVERRNYTLKLRREDIDFLRHINVTIN 1318
DB 1259 VAVRLKESYSTEKBAWQIOETAPPSGWPQVGYVERRNYTLKLRREDIDFLRHINVTIN 1318
QY 1319 GGEKVGIVGRGAKSSLTGLFRINSAGEEIIIDGINAKIGLHDLRFKITTIPDPV 1378
DB 1319 GGEKVGIVGRGAKSSLTGLFRINSAGEEIIIDGINAKIGLHDLRFKITTIPDPV 1378
QY 1379 LFGSGLRNLDLPFGQYSDDEEYVTSLELAHLKDQVYSALPDKLHDCAGEGENLSVQORQV 1438
DB 1379 LFGSGLRNLDLPFGQYSDDEEYVTSLELAHLKDQVYSALPDKLHDCAGEGENLSVQORQV 1438
QY 1436 CLARALRKTKILVIDEATAVDLETDNLIOSTIRTOGEDCTVLTIAHRLNTINDYTVI 1495
DB 1436 CLARALRKTKILVIDEATAVDLETDNLIOSTIRTOGEDCTVLTIAHRLNTINDYTVI 1495

RESULT 15
US-08-462-109A-6
Sequence 6, Application US/08462109A
Patent No. 5882875
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,109A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-6

Query Match 70.7%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALGFGCAGSDSDPLMDNNTWNTSNPTFCFQNTLVVWVPCFYLACFPFFLYLSRH 60
DB 1 MALGFGCAGSDSDPLMDNNTWNTSNPTFCFQNTLVVWVPCFYLACFPFFLYLSRH 60
QY 61 DRGVIOMTPLNKTALGFLMIYCMADLFYSFMRSGRIFLAPFVLSPTLLGTTLLA 120
DB 61 DRGVIOMTPLNKTALGFLMIYCMADLFYSFMRSGRIFLAPFVLSPTLLGTTLLA 120
QY 121 TFLIQLEBRKGVQSSGIMLTFMVALVCALAILRSKIMTALKEDAQVDLFFRDIIFYVFS 180
DB 121 TFLIQLEBRKGVQSSGIMLTFMVALVCALAILRSKIMTALKEDAQVDLFFRDIIFYVFS 180
QY 181 LLLIQVLVLSGFSRSPFSETIHDNPNCPSSSASFLSKITITWMTGLIVRGYROPLEGSD 240
DB 181 LLLIQVLVLSGFSRSPFSETIHDNPNCPSSSASFLSKITITWMTGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSQVYVPLVKNMKKECAKTRKQPVKVYVS-SKDPNQPRESSKVDNNEVEEA 299
DB 241 LMSLNKEDTSQVYVPLVKNMKKECAKTRKQPVKVYVS-SKDPNQPRESSKVDNNEVEEA 299
QY 300 LIVKSPKEMNPSEFLVLYKTEGYPFLMSFFKAIHDLMPFGPQILKLILKFNVDYKAP 359
DB 300 LIVKSPKEMNPSEFLVLYKTEGYPFLMSFFKAIHDLMPFGPQILKLILKFNVDYKAP 359
QY 360 DMQGYFTTVLLFVPAQCLQTLVHQYFHI CFYSGMRITKAVT GAAYRKALVITNSARISST 419
DB 360 DMQGYFTTVLLFVPAQCLQTLVHQYFHI CFYSGMRITKAVT GAAYRKALVITNSARISST 419
QY 420 VGEIVNLSVDAQGFMDLATYINNIWASAPLOVIALYLLMTNLGPSYLAGAVAVWLVNVP 479
DB 420 VGEIVNLSVDAQGFMDLATYINNIWASAPLOVIALYLLMTNLGPSYLAGAVAVWLVNVP 479
QY 480 NAYVAMKTKTYOVAVHMSKDNRIKLMNEILNGIVLKL YAMELAFKQVLAIRQEBELKVL 539
DB 480 NAYVAMKTKTYOVAVHMSKDNRIKLMNEILNGIVLKL YAMELAFKQVLAIRQEBELKVL 539

QY 540 KKSAYLSAVGFTFWCTPELVALCTPAVYVTTIDENNIIIDAQTAFLALFNILREPLNII 599
DB 541 KKSAYLSAVGFTFWCTPELVALCTPAVYVTTIDENNIIIDAQTAFLALFNILREPLNII 600
QY 600 PMVISTIVQASVSLKRLRIPLSHELEBDSIERRPVKDGGTNSITVANAFTWASDDP 659
DB 601 PMVISTIVQASVSLKRLRIPLSHELEBDSIERRSIKSGEG-NGITVGNALFTWASGPP 659
QY 660 TLNGITPSIPGALVAVVGVGCKSSLSLALAEMLDMKVEGHVAIKSSVAVVPOQAWION 719
DB 660 TLNGITPSIPGALVAVVGVGCKSSLSLALAEMLDMKVEGHVAIKSSVAVVPOQAWION 719
QY 720 DSIARENILFGQLEBPYRASYIQCALLPDLEIIPSGDRTEIGKGVNLSSGQORVSLA 779
DB 720 DSIARENILFGQLEBPYRASYIQCALLPDLEIIPSGDRTEIGKGVNLSSGQORVSLA 779
QY 780 RAYVSNADIVLPDPLSAVDAAVGHKHFENVIGPKMLKNTRIILVTHSMGYLPQVDYI 839
DB 780 RAYVSNADIVLPDPLSAVDAAVGHKHFENVIGPKMLKNTRIILVTHSMGYLPQVDYI 839
QY 840 VMSGKISSEMGSYOELLARDGAFAPLRTYASTBOEDAEENGVTGVSGPEKAKOMENG 899
DB 840 VMSGKISSEMGSYOELLARDGAFAPLRTYASTBOEDAEENGVTGVSGPEKAKOMENG 899
QY 900 MLVYDSAGKOLQROLSSSSSYSGBISRHNSHSTAIOQAABKEKERTWLMENDAKYQGVK 959
DB 900 MLVYDSAGKOLQROLSSSSSYSGBISRHNSHSTAIOQAABKEKERTWLMENDAKYQGVK 959
QY 960 LSVYMDYMKAIQGLFISFISFLPMCNHVSALASNYMLSLWTD-PYNGTOEHKVRSLV 1018
DB 956 LSVYMDYMKAIQGLFISFISFLPMCNHVSALASNYMLSLWTD-PYNGTOEHKVRSLV 1015
QY 1019 YGALGISQGIAPFGYSNAVSIGIILASRCLHVDLHLSILSPMSFERTPSGNLVNRESK 1078
DB 1016 YGALGISQGIAPFGYSNAVSIGIILASRCLHVDLHLSILSPMSFERTPSGNLVNRESK 1075
QY 1079 ELDTVDMSIPEVIMFNGSLFNVIGACIVILLATPIAIIIPGLIYFFQRYVASSR 1138
DB 1076 ELDTVDMSIPEVIMFNGSLFNVIGACIVILLATPIAIIIPGLIYFFQRYVASSR 1135
QY 1139 QLRLEBSVRSPPVSHNETLLGVSVIRAFEBORFIHOSDLKYDENOKAYYPSIVANRW 1198
DB 1136 QLRLEBSVRSPPVSHNETLLGVSVIRAFEBORFIHOSDLKYDENOKAYYPSIVANRW 1195
QY 1199 LAVRLKCVGNCIVLFAALFAVISRHSLSAGLVGSYSLOVTTYLNMLVRMSSEMETNI 1258
DB 1196 LAVRLKCVGNCIVLFAALFAVISRHSLSAGLVGSYSLOVTTYLNMLVRMSSEMETNI 1255
QY 1259 VAVERLKEYSETERKAAWQIOETNAPBSMPOVGRVVERFRNYCLARYEDLDVYLRHINTIN 1318
DB 1256 VAVERLKEYSETERKAAWQIOETNAPBSMPOVGRVVERFRNYCLARYEDLDVYLRHINTIN 1315
QY 1319 GGEKVGIIVGRTGAGKSLTGLPRINSAGEEIIIDGINIAKGLHDLRFKTIITIPDPV 1378
DB 1316 GGEKVGIIVGRTGAGKSLTGLPRINSAGEEIIIDGINIAKGLHDLRFKTIITIPDPV 1375
QY 1379 LFGSGSLRMNLDPSQYSDBEWMTSLBLAHLKDFVSAJLBDLDEHCAEGENLSYQGRQV 1438
DB 1376 LFGSGSLRMNLDPSQYSDBEWMTSLBLAHLKDFVSAJLBDLDEHCAEGENLSYQGRQV 1435
QY 1439 CLARALIRKTKIIVLDEATAVAVLETDLLQSTIRIROPEDCTVLTAAHRLNTIMDYTRVI 1498
DB 1436 CLARALIRKTKIIVLDEATAVAVLETDLLQSTIRIROPEDCTVLTAAHRLNTIMDYTRVI 1495
QY 1499 VLDKGEIOEYGAPEDDLQQRGLFFYSMAKDAGLV 1531
DB 1496 VLDKGEIOEYGAPEDDLQQRGLFFYSMAKDAGLV 1528

Search completed: December 15, 2005, 15:49:51
Job time : 48.0811 secs

Blank (uspo)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:00 ; Search time 159.832 Seconds
(without alignments)
5037.528 Million cell updates/sec

Title: US-10-665-283-6

Perfect score: 9901

Sequence: 1 MALRGFCSADGSDPLMDMNV.....ASVAVAKAKKXFSISPSLS 1927

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9901	100.0	1927	US-10-665-283-6	Sequence 6, Appli
2	9891	99.9	1927	US-10-665-283-1	Sequence 1, Appli
3	9861	99.6	1947	US-10-665-283-8	Sequence 8, Appli
4	9734	98.3	1891	US-10-665-283-4	Sequence 4, Appli
5	7860	79.4	1531	US-09-939-853A-86	Sequence 86, Appli
6	7860	79.4	1531	US-10-667-891-6	Sequence 6, Appli
7	7860	79.4	1531	US-10-807-466-6	Sequence 6, Appli
8	7860	79.4	1531	US-10-484-577-678	Sequence 678, App
9	7860	79.4	1531	US-10-889-503-19	Sequence 19, Appli
10	7860	79.4	1531	US-10-756-149-5033	Sequence 5033, Ap
11	7769	78.5	1515	US-09-939-853A-87	Sequence 87, Appli
12	7763	78.4	1515	US-10-618-281-42	Sequence 42, Appli
13	7037	71.1	1388	US-10-408-765A-1718	Sequence 1718, Ap
14	7002.5	70.7	1528	US-09-939-853A-88	Sequence 88, Appli
15	6068.5	61.3	1303	US-10-618-281-63	Sequence 63, Appli
16	4854.5	49.0	1911	US-10-665-283-2	Sequence 2, Appli
17	4637.5	47.4	1875	US-10-665-283-5	Sequence 5, Appli
18	4667.5	47.1	1895	US-10-665-283-9	Sequence 9, Appli
19	4467.5	45.4	1901	US-10-665-283-10	Sequence 10, Appli
20	4491.5	45.4	1527	US-09-939-853A-83	Sequence 83, Appli
21	4491.5	45.4	1527	US-09-939-853A-84	Sequence 84, Appli
22	4491.5	45.4	1527	US-10-235-027-1342	Sequence 1342, Ap
23	4487.5	45.3	1527	US-10-889-503-5	Sequence 5, Appli
24	4487.5	45.3	1530	US-10-889-503-3	Sequence 3, Appli
25	4479.5	44.4	1530	US-10-618-281-55	Sequence 55, Appli
26	4396	44.4	1522	US-09-939-853A-85	Sequence 85, Appli
27	4123	41.6	1941	US-10-665-283-12	Sequence 12, Appli

28	4093	41.3	1961	5	US-10-665-283-7	Sequence 7, Appli
29	3966	40.1	1905	5	US-10-665-283-3	Sequence 3, Appli
30	3879.5	39.2	1548	4	US-10-807-466-28	Sequence 28, Appli
31	3860.5	39.0	1548	4	US-10-807-466-1	Sequence 1, Appli
32	3860.5	39.0	1548	5	US-10-807-466-1	Sequence 1, Appli
33	3846	38.8	1538	5	US-10-807-466-23	Sequence 23, Appli
34	3841	38.8	1538	5	US-10-807-466-25	Sequence 25, Appli
35	3835	38.7	1538	5	US-10-807-466-24	Sequence 24, Appli
36	3817.5	38.6	1539	5	US-10-807-466-26	Sequence 26, Appli
37	3817	38.6	1591	6	US-11-097-143-13266	Sequence 13266, A
38	3797	38.3	1538	5	US-10-807-466-27	Sequence 27, Appli
39	3779	38.2	1538	5	US-10-807-466-16	Sequence 16, Appli
40	3774	38.1	1538	5	US-10-807-466-18	Sequence 18, Appli
41	3773	38.1	1538	5	US-10-807-466-22	Sequence 22, Appli
42	3768	38.1	1538	5	US-10-807-466-17	Sequence 17, Appli
43	3754	37.9	1538	5	US-10-807-466-19	Sequence 19, Appli
44	3750.5	37.9	1539	5	US-10-807-466-20	Sequence 20, Appli
45	3733	37.7	1547	6	US-11-097-143-25248	Sequence 25248, A

ALIGNMENTS

RESULT 1									
US-10-665-283-6									
Sequence 6, Application US/10665283									
Publication No. US20050063989A1									
GENERAL INFORMATION:									
APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE									
APPLICANT: DERAND, Renaud									
APPLICANT: GARCIA, Elisabeth									
APPLICANT: PROST, Anne-lise									
APPLICANT: REYTLAUD, Jean									
APPLICANT: VIVAUDOU, Michel									
TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREO									
FILE REFERENCE: P263US104									
CURRENT APPLICATION NUMBER: US/10/665,283									
CURRENT FILING DATE: 2003-09-22									
NUMBER OF SEQ ID NOS: 25									
SOFTWARE: Patentin version 3.1									
SEQ ID NO 6									
LENGTH: 1927									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-10-665-283-6									
Query Match									
Best Local Similarity 100.0%; Score 9901; DB 5; Length 1927;									
Matches 1927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MALRGFCSADGSDPLMDMNTWNTSNDFTKCFONTLVVWPCFYLMACFPFYLYSRH	60						
DB	1	MALRGFCSADGSDPLMDMNTWNTSNDFTKCFONTLVVWPCFYLMACFPFYLYSRH	60						
QY	61	DRGTYOMTPNKTKTALGFLMTVCADLTFYFWEBSRGFLAPVPVYSLTGITTLA	120						
DB	61	DRGTYOMTPNKTKTALGFLMTVCADLTFYFWEBSRGFLAPVPVYSLTGITTLA	120						
QY	121	TFPIQLERRRGVSGGIMLTFMLVALYCALAIRSKIMTLAKSDAQVDLFRDITFYVFS	180						
DB	121	TFPIQLERRRGVSGGIMLTFMLVALYCALAIRSKIMTLAKSDAQVDLFRDITFYVFS	180						
QY	181	LLLIQLVLSGFSRSPFLFSETIHDPNFCPSASFLSRITFWMITGLIVGYQPLEGSD	240						
DB	181	LLLIQLVLSGFSRSPFLFSETIHDPNFCPSASFLSRITFWMITGLIVGYQPLEGSD	240						
QY	241	LMSLNKEDSEGVVPTVVKWKKECAKTRQPKVYVSSDPAOPKSSSVYDNEVEAL	300						
DB	241	LMSLNKEDSEGVVPTVVKWKKECAKTRQPKVYVSSDPAOPKSSSVYDNEVEAL	300						
QY	301	IVKSPQKKNPISLFKVLKTFGFPFLMSFFFKAIHDLMTSGQIIKLTKLFVNDTKAPD	360						
DB	301	IVKSPQKKNPISLFKVLKTFGFPFLMSFFFKAIHDLMTSGQIIKLTKLFVNDTKAPD	360						

QY 361 WQGFYTLVLTAVTACLTQLVTHQYHICFVSGMRIKTAVIGAVYKALVTINSARKSSTV 420
DB 361 WQGFYTLVLTAVTACLTQLVTHQYHICFVSGMRIKTAVIGAVYKALVTINSARKSSTV 420
QY 421 GEIYNLMSVDNORFMDLATYINMISAPLOVTLAYLMLNIGBSVLAGVAVMTLVN 480
DB 421 GEIYNLMSVDNORFMDLATYINMISAPLOVTLAYLMLNIGBSVLAGVAVMTLVN 480
QY 481 AVAMAKTKTYOVAMHKSNDRIKLMNEILNGIKVLKLYAMELAFDXYLAIRQELKYLK 540
DB 481 AVAMAKTKTYOVAMHKSNDRIKLMNEILNGIKVLKLYAMELAFDXYLAIRQELKYLK 540
QY 541 KSAVLSAVGTTTWCCTPPLVALCTPAAVYVTIDENNIIDAQTAFFVSALFNILRPPLNITL 600
DB 541 KSAVLSAVGTTTWCCTPPLVALCTPAAVYVTIDENNIIDAQTAFFVSALFNILRPPLNITL 600
QY 601 MVISSIVASVSLKRLIFLSHEELEPDSIERRPVKDGCTNSITVRNATFTWASDPT 660
DB 601 MVISSIVASVSLKRLIFLSHEELEPDSIERRPVKDGCTNSITVRNATFTWASDPT 660
QY 661 LNGITFSIPEGALVAVVGVQCGKSSLLSALLAEMDKYGHVAIKGSVAYVPOQAMIOND 720
DB 661 LNGITFSIPEGALVAVVGVQCGKSSLLSALLAEMDKYGHVAIKGSVAYVPOQAMIOND 720
QY 721 SLRENIIFGCQLEBPYRSVIOACALPDLBITLPSGDRTEIGEKVNLSCGQKQKRSIAR 780
DB 721 SLRENIIFGCQLEBPYRSVIOACALPDLBITLPSGDRTEIGEKVNLSCGQKQKRSIAR 780
QY 781 AVYNADITYLEDDPLSAADAVGKHI FENVI GPKMKLNKTRILVTHSMSTLPQVDVITV 840
DB 781 AVYNADITYLEDDPLSAADAVGKHI FENVI GPKMKLNKTRILVTHSMSTLPQVDVITV 840
QY 841 MSGGKISMSGYQELIARDGAFABEFLRTYASTEQDAEENGVTGVSQPKAKOMENGM 900
DB 841 MSGGKISMSGYQELIARDGAFABEFLRTYASTEQDAEENGVTGVSQPKAKOMENGM 900
QY 901 LVTDSAGKQOLRSSSSSSYSGDISRHNSTAELOKAKKEETWKLMEADKQGYKL 960
DB 901 LVTDSAGKQOLRSSSSSSYSGDISRHNSTAELOKAKKEETWKLMEADKQGYKL 960
QY 961 SVYNDYMAKILFTSFLSIFLPMCHNVALSNVLSLMTDDPIYNGOETHKURLSYG 1020
DB 961 SVYNDYMAKILFTSFLSIFLPMCHNVALSNVLSLMTDDPIYNGOETHKURLSYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGGILASRCHVDLHSLRSPMSFPERTPSGMLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGGILASRCHVDLHSLRSPMSFPERTPSGMLVNRFSKEL 1080
QY 1081 DTVDSMIPEVIKMGSLFENVI GACIVILATPIAIIIPPLGIYFFVQRPYVASSRQL 1140
DB 1081 DTVDSMIPEVIKMGSLFENVI GACIVILATPIAIIIPPLGIYFFVQRPYVASSRQL 1140
QY 1141 KRLESVRSRVSYSHPNEITLGVSVIRAFEBQERFHODKLVKDEOKAYYSIVANRWLA 1200
DB 1141 KRLESVRSRVSYSHPNEITLGVSVIRAFEBQERFHODKLVKDEOKAYYSIVANRWLA 1200
QY 1201 VRLECVGNCIYLPALFAVISRHSLSAGLVGYSYSLOVTTYLWMLVRMSSEMETNIVA 1260
DB 1201 VRLECVGNCIYLPALFAVISRHSLSAGLVGYSYSLOVTTYLWMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQGRVEFRNCTARYREDDLVLAHINVTINGG 1320
DB 1261 VERLKEYSETEKAPWQIOETAPSSWPQGRVEFRNCTARYREDDLVLAHINVTINGG 1320
QY 1321 EKVGIVGRTGAKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
DB 1321 EKVGIVGRTGAKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
QY 1381 SSGSRMNLDPFSQYSDDEEVTWLSLELAHKDFVSALPDLDEHCAAGGNLSVGOQOLVCL 1440
DB 1381 SSGSRMNLDPFSQYSDDEEVTWLSLELAHKDFVSALPDLDEHCAAGGNLSVGOQOLVCL 1440

QY 1441 APALIRKTKIIVDEBATAVDLEFDLLOSTIRFOQFEDCTVLTAAHRLNTIMDYTRIVL 1500
DB 1441 APALIRKTKIIVDEBATAVDLEFDLLOSTIRFOQFEDCTVLTAAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGAUSDLLQORGLFYSMAKAGLVGGGGGKLSKGIIPERYVLTALAEPAE 1560
DB 1501 DKGEIOEYGAUSDLLQORGLFYSMAKAGLVGGGGGKLSKGIIPERYVLTALAEPAE 1560
QY 1561 PRYTRERRARFVSKGNCVNAHKNIREQGRFLODVFTTLVLDLXPHLILFTMSFLCSW 1620
DB 1561 PRYTRERRARFVSKGNCVNAHKNIREQGRFLODVFTTLVLDLXPHLILFTMSFLCSW 1620
QY 1621 LLPAMWMLIAFAHGDIAPEGSTNVCVTSIHSFSSAFLSIEVQVTIGFGGRWTEBCP 1680
DB 1621 LLPAMWMLIAFAHGDIAPEGSTNVCVTSIHSFSSAFLSIEVQVTIGFGGRWTEBCP 1680
QY 1681 LAILILIVQNVGMINAMIGCI FMTQAQHRBAETLIFSKHAVITLRHRLCPMLRPG 1740
DB 1681 LAILILIVQNVGMINAMIGCI FMTQAQHRBAETLIFSKHAVITLRHRLCPMLRPG 1740
QY 1741 DLKRSMTIISATIHQVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAFLIIVHIDS 1800
DB 1741 DLKRSMTIISATIHQVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAFLIIVHIDS 1800
QY 1801 NSPLYDLAPSDLHHODLEIIVILEGVETTGITTOARTSYLADEILMGQFVPIVAED 1860
DB 1801 NSPLYDLAPSDLHHODLEIIVILEGVETTGITTOARTSYLADEILMGQFVPIVAED 1860
QY 1861 GRYSVDYSKFGNTIKVPPICLTARQDEBDSLALTLASRGPLRAASVAAKAPRES 1920
DB 1861 GRYSVDYSKFGNTIKVPPICLTARQDEBDSLALTLASRGPLRAASVAAKAPRES 1920
QY 1921 ISPDSL 1927
DB 1921 ISPDSL 1927

RESULT 2
US-10-665-283-1
; Sequence 1, Application US/10665283
; Publication No. US2005063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLAUD, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: P263US104
; CURRENT APPLICATION NUMBER: US/10/665,283
; CURRENT FILING DATE: 2003-09-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1927
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-665-283-1

Query Match 99.9%; Score 9891; DB 5; Length 1927;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTWTNTSNDFTKCFONTYLVWVPCYYLMAACEPFYFLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTWTNTSNDFTKCFONTYLVWVPCYYLMAACEPFYFLYSRH 60
QY 61 DRGIQMTPLNKTKTALGFLIMIVCMADLFPFSFERSRGFLAIVPFLVSPFLIGITTTLA 120
DB 61 DRGIQMTPLNKTKTALGFLIMIVCMADLFPFSFERSRGFLAIVPFLVSPFLIGITTTLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180

```

Db      121  TFLQLRRKRGVSSGIMLTFMLVALVCAALILRSKIMTALKEDAOVDFDIIFYYFS 180
Qy      181  LLLIQLVLSGSDSPLFSETIHDNCPRESSASPLSITFWMTGLIVRGYRPLGSD 240
Db      181  LLLIQLVLSGSDSPLFSETIHDNCPRESSASPLSITFWMTGLIVRGYRPLGSD 240
Qy      241  LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAPKSSSKVDANEVEAL 300
Db      241  LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAPKSSSKVDANEVEAL 300
Qy      301  IVSPQKEMNPSELFKVLYKTFGPFILMSFFPKAIIHDLMMFSGPOLIKLLIFVNDTRAPD 360
Db      301  IVSPQKEMNPSELFKVLYKTFGPFILMSFFPKAIIHDLMMFSGPOLIKLLIFVNDTRAPD 360
Qy      361  WOGFFYVLLFVTRACIQLVLHQYFHCIFVSGMFKTAIVGAVYRKALVTNSARKSSTV 420
Db      361  WOGFFYVLLFVTRACIQLVLHQYFHCIFVSGMFKTAIVGAVYRKALVTNSARKSSTV 420
Qy      421  GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLMNLGSPVLGAVAVMTLAPVN 480
Db      421  GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLMNLGSPVLGAVAVMTLAPVN 480
Qy      481  AVNAMKTKTYQVAHMKSKDNRIKLMNEILINGIKVLKUYAMELAFKDVLAIROBELVVK 540
Db      481  AVNAMKTKTYQVAHMKSKDNRIKLMNEILINGIKVLKUYAMELAFKDVLAIROBELVVK 540
Qy      541  KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFSALFNLIRPLILP 600
Db      541  KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFSALFNLIRPLILP 600
Qy      601  MWISSIQAASVLSKRLRIFLSHEELBPSIERPVKGGGNSITVANAFTWARSNDPT 660
Db      601  MWISSIQAASVLSKRLRIFLSHEELBPSIERPVKGGGNSITVANAFTWARSNDPT 660
Qy      661  LNCITFESIPGALVAVVGVCCKSSLSALLAEMDVKEGVAIKGSVAVYPOQAWIOND 720
Db      661  LNCITFESIPGALVAVVGVCCKSSLSALLAEMDVKEGVAIKGSVAVYPOQAWIOND 720
Qy      721  SLEENILFGCOLBEPYRSYIOACALLPDLIELPSGRTIEGKGVNLSSGQOKORVSLAR 780
Db      721  SLEENILFGCOLBEPYRSYIOACALLPDLIELPSGRTIEGKGVNLSSGQOKORVSLAR 780
Qy      781  AAVSNADIVLPDPLSAVDAHVGHIFENYIGPKMKKNTRIILYHSMGYLPOVDYIIV 840
Db      781  AAVSNADIVLPDPLSAVDAHVGHIFENYIGPKMKKNTRIILYHSMGYLPOVDYIIV 840
Qy      841  MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOEQDAEENGVTGVSQPKKAKOMENGM 900
Db      841  MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOEQDAEENGVTGVSQPKKAKOMENGM 900
Qy      901  LVYDSAGKQOLQROLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKMEADKAQTQVKL 960
Db      901  LVYDSAGKQOLQROLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKMEADKAQTQVKL 960
Qy      961  SVYMDTWKALGLTSLSTFLFMCNHYSAASNYWLSLWDDPVLNGTQHTKVRISVYG 1020
Db      961  SVYMDTWKALGLTSLSTFLFMCNHYSAASNYWLSLWDDPVLNGTQHTKVRISVYG 1020
Qy      1021  ALGISOGIAVFGYSMAVISIGIILASRCLAYDLHLSIRSPMSFEERTPSGNLVNRSKEL 1080
Db      1021  ALGISOGIAVFGYSMAVISIGIILASRCLAYDLHLSIRSPMSFEERTPSGNLVNRSKEL 1080
Qy      1081  DTYDSMIPVYIKMFMSLFFNVIGACIVILATPAAIIPPLGLIYFVQRFVASSRCL 1140
Db      1081  DTYDSMIPVYIKMFMSLFFNVIGACIVILATPAAIIPPLGLIYFVQRFVASSRCL 1140
Qy      1141  KRLSEYSRSPVYSHFNETLLGVSVIRAFEEOEERFIHOSDLKVDENQAAVYPSIVANRWLA 1200
Db      1141  KRLSEYSRSPVYSHFNETLLGVSVIRAFEEOEERFIHOSDLKVDENQAAVYPSIVANRWLA 1200
Qy      1201  VRLECVGNCLVFAALFAVISRHSLSAGLVGLSVSYSIQVTTYLNMVLVRMSSEMETNIVA 1260

```

```

Db      1201  VRLECVGNCLVFAALFAVISRHSLSAGLVGLSVSYSIQVTTYLNMVLVRMSSEMETNIVA 1260
Qy      1261  VERLEKESYEBKAPMIOIOSTAPSSWPQVGRVFRNYCYLRREDLDPYLRHINVTNGS 1320
Db      1261  VERLEKESYEBKAPMIOIOSTAPSSWPQVGRVFRNYCYLRREDLDPYLRHINVTNGS 1320
Qy      1321  EKVIGVGRTAGKSSSLTGLFPRINESAGEIILIDGINIAKIGLHDLAFKIIIPQDPVLF 1380
Db      1321  EKVIGVGRTAGKSSSLTGLFPRINESAGEIILIDGINIAKIGLHDLAFKIIIPQDPVLF 1380
Qy      1381  SGLIRNMLDPSQYSDSEEWTSLELAHLKDFVSLAPDKLHCEAEGENISVGOQOLVCL 1440
Db      1381  SGLIRNMLDPSQYSDSEEWTSLELAHLKDFVSLAPDKLHCEAEGENISVGOQOLVCL 1440
Qy      1441  ARALRKTILVNDDEATAVADLETDLLIOSTIRQFQFEDCVTLTAHRLNTIMOTRYIVL 1500
Db      1441  ARALRKTILVNDDEATAVADLETDLLIOSTIRQFQFEDCVTLTAHRLNTIMOTRYIVL 1500
Qy      1501  DKGEIOEYGAPSDLOOQGLFYSMADKAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Db      1501  DKGEIOEYGAPSDLOOQGLFYSMADKAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Qy      1561  PRYTRERRARFVSKKNCNVAAKNIREQGRFLQDVFTTLVDLKWPHTLIIFTWSFLCSW 1620
Db      1561  PRYTRERRARFVSKKNCNVAAKNIREQGRFLQDVFTTLVDLKWPHTLIIFTWSFLCSW 1620
Qy      1621  LFLPAMWMLIAFAHGLDAPBEGTNVPCVTSIHSPSSAFLEISIEVOTTIGGGRMTECP 1680
Db      1621  LFLPAMWMLIAFAHGLDAPBEGTNVPCVTSIHSPSSAFLEISIEVOTTIGGGRMTECP 1680
Qy      1681  LAILILIVONIYGIMINAMLGCI FMKTAQAHRAETLISKHAVITLHRGRICFMRVVG 1740
Db      1681  LAILILIVONIYGIMINAMLGCI FMKTAQAHRAETLISKHAVITLHRGRICFMRVVG 1740
Qy      1741  DLKSMISATIMHMOVVKRTSPEGEVPLHQVDIPWENGVGNGIFLVAPLIIVHYDS 1800
Db      1741  DLKSMISATIMHMOVVKRTSPEGEVPLHQVDIPWENGVGNGIFLVAPLIIVHYDS 1800
Qy      1801  NSPLYDLAPSDLHHODLEIIVLEGEVETTGITTOARTSYLADEIIMGQRFVPIVAED 1860
Db      1801  NSPLYDLAPSDLHHODLEIIVLEGEVETTGITTOARTSYLADEIIMGQRFVPIVAED 1860
Qy      1861  GRYSVDYSKRGNTIKVPTPLCTAROLDDEDSLDDALLTLASSRQPLRGRSVAAKAPKFS 1920
Db      1921  ISPDLS 1927

```

```

RESULT 3
US-10-665-283-8
; Sequence 8, Application US/10665283
; Publication No. US20050063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLON, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILER REFERENCE: F263US104
; CURRENT APPLICATION NUMBER: US/10/665,283
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1947
; TYPE: PRK
; ORGANISM: Homo sapiens
US-10-665-283-8

```

Query Match 99.6%; Score 9861; DB 5; Length 1947;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1925; Conservative 0; Mismatches 2; Indels 20; Gaps 2;

QY 1 MALRGFSGADSDLDLMDMNTWNTSNPDFTKCFONTVLVWPCFYLMACFPFFLYLSRH 60
 DB 1 MALRGFSGADSDLDLMDMNTWNTSNPDFTKCFONTVLVWPCFYLMACFPFFLYLSRH 60
 QY 61 DRGIQMTPLNKTKTALGELMTICMDLFFSFMERSRGIFLAPFVLSPTLLGITLLA 120
 DB 61 DRGIQMTPLNKTKTALGELMTICMDLFFSFMERSRGIFLAPFVLSPTLLGITLLA 120
 QY 121 TFLIOLERRKGVSSGIMLTFMLVALCALAILRSKIMTALKEDAQVDLFFDITFYVYS 180
 DB 121 TFLIOLERRKGVSSGIMLTFMLVALCALAILRSKIMTALKEDAQVDLFFDITFYVYS 180
 QY 181 LLLIQLVLSGFSPLFSETIHDNPPCESSASFLSITFWITGLIVRGYROPLEBSD 240
 DB 181 LLLIQLVLSGFSPLFSETIHDNPPCESSASFLSITFWITGLIVRGYROPLEBSD 240
 QY 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKOPVKVYSSKDPADPKSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKOPVKVYSSKDPADPKSSKVDANEVEAL 300
 QY 301 IVKSPOKEMNBSLFKVLYKTFGPYFLMSFFKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
 DB 301 IVKSPOKEMNBSLFKVLYKTFGPYFLMSFFKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
 QY 361 WQGFYTVLLFVTAQLOTLVHOYFHICFVSGMRITKAVIGAVYKALVITNSARKSTV 420
 DB 361 WQGFYTVLLFVTAQLOTLVHOYFHICFVSGMRITKAVIGAVYKALVITNSARKSTV 420
 QY 421 GEIYNIMSVDORFMDLTYINMISAPLOYLIALYLLMLNGPSVLGAVMVLVNVN 480
 DB 421 GEIYNIMSVDORFMDLTYINMISAPLOYLIALYLLMLNGPSVLGAVMVLVNVN 480
 QY 481 AVMAAMTKTYOVANHKSKDNRIKLNNEILNGIKVLKLYAMELAFKDKYALROBELKYLK 540
 DB 481 AVMAAMTKTYOVANHKSKDNRIKLNNEILNGIKVLKLYAMELAFKDKYALROBELKYLK 540
 QY 541 KSAVLSAVGTWCTPFLVALCTPAVYVTTDENNILDQAOTAFVSLAFNLLRPLNTLP 600
 DB 541 KSAVLSAVGTWCTPFLVALCTPAVYVTTDENNILDQAOTAFVSLAFNLLRPLNTLP 600
 QY 601 MVISIVASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
 DB 601 MVISIVASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
 QY 661 LMGITFSPREGALVAVVGVCGKSSLSALLAEMDKYEGHVALKGSVAVYVPOQAMIOND 720
 DB 661 LMGITFSPREGALVAVVGVCGKSSLSALLAEMDKYEGHVALKGSVAVYVPOQAMIOND 720
 QY 721 SLRENIILFCQLEEBYRVSIVQACALLPDLLETLPSSGDRTEIGEKVNLSSGQKORVSLAR 780
 DB 721 SLRENIILFCQLEEBYRVSIVQACALLPDLLETLPSSGDRTEIGEKVNLSSGQKORVSLAR 780
 QY 781 AVYSNADITYLPDDPLSAVDAAVGHKIFENNVIGPKMKLNKTRILVTHSMSTYLPQVDVIV 840
 DB 781 AVYSNADITYLPDDPLSAVDAAVGHKIFENNVIGPKMKLNKTRILVTHSMSTYLPQVDVIV 840
 QY 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOQDAEENGVTGVSGRGKAKEMNGM 900
 DB 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOQDAEENGVTGVSGRGKAKEMNGM 900
 QY 901 LVTDASAGQOLQROUSSSSYSGDISRHNSSTAELQAEAKKETWKLMEADKAQOTGOYL 960
 DB 901 LVTDASAGQOLQROUSSSSYSGDISRHNSSTAELQAEAKKETWKLMEADKAQOTGOYL 960
 QY 961 SVYMDYMAIGLIFLSIFLFLPMCHVSAALASNYLSLMTDPIYNGVOEHTKVALSYG 1020
 DB 961 SVYMDYMAIGLIFLSIFLFLPMCHVSAALASNYLSLMTDPIYNGVOEHTKVALSYG 1020

QY 1021 ALGISQIAVFGYSMAVSGIILASRCLHVDLHSILRSFMSFPERTPSGNLVNRBEKEL 1080
 DB 1021 ALGISQIAVFGYSMAVSGIILASRCLHVDLHSILRSFMSFPERTPSGNLVNRBEKEL 1080
 QY 1081 DTVDSMTPEVITKPMGSLFENVIGACTIYLATPAAIIPPLGIYFVQRFVYASSROL 1140
 DB 1081 DTVDSMTPEVITKPMGSLFENVIGACTIYLATPAAIIPPLGIYFVQRFVYASSROL 1140
 QY 1141 KRLSVSRSPVSHFNTELLGVSVIRAFEEQERFIHOSDKVDENQAKYPSIVANWMLA 1200
 DB 1141 KRLSVSRSPVSHFNTELLGVSVIRAFEEQERFIHOSDKVDENQAKYPSIVANWMLA 1200
 QY 1201 VRLCEVNCIYLPALPAVIRSRHSAGLVGLSVSYLQVTTYLNMVLVMSSEMETNIVA 1260
 DB 1201 VRLCEVNCIYLPALPAVIRSRHSAGLVGLSVSYLQVTTYLNMVLVMSSEMETNIVA 1260
 QY 1261 VERLKEYSETEKAPMOIOTAPSSWPQVGRVFRVYCLRYRBDLPVLRHINVTNGG 1320
 DB 1261 VERLKEYSETEKAPMOIOTAPSSWPQVGRVFRVYCLRYRBDLPVLRHINVTNGG 1320
 QY 1321 EKVGI VGRGTAGKSLTGLFRINESAGEIIDIIGINIAKIGLHDLRFKTIIPQDPVL 1380
 DB 1321 EKVGI VGRGTAGKSLTGLFRINESAGEIIDIIGINIAKIGLHDLRFKTIIPQDPVL 1380
 QY 1381 SGLRNMULDPPSOYSDDEEWTSLLEALHKDFVSALPKLDHECAEGENLSVGROLVCL 1440
 DB 1381 SGLRNMULDPPSOYSDDEEWTSLLEALHKDFVSALPKLDHECAEGENLSVGROLVCL 1440
 QY 1441 ARALLRTKILVDEAFAANDLETDDLIQSTIRQPFEDCVTLTAHRLNTMDTRVYL 1500
 DB 1441 ARALLRTKILVDEAFAANDLETDDLIQSTIRQPFEDCVTLTAHRLNTMDTRVYL 1500
 QY 1501 DKGBIOEYGAPSDLOQRGLFYSMADAGLVGGGGGMLSRKGIIPPEYVLTRLAEDPAE 1560
 DB 1501 DKGBIOEYGAPSDLOQRGLFYSMADAGLVGGGGGMLSRKGIIPPEYVLTRLAEDPAE 1560
 QY 1561 PRYTRERRARFVSKKNCNVAAHKNIREQGFLODFTTVLDLKWPHULLIFWSPICSW 1620
 DB 1561 PRYTRERRARFVSKKNCNVAAHKNIREQGFLODFTTVLDLKWPHULLIFWSPICSW 1620
 QY 1621 LLFAMWMLLAFAG-----DLAP-----GBGNVPCVTSIHSSFSALF 1680
 DB 1621 LLFAMWMLLAFAG-----DLAP-----GBGNVPCVTSIHSSFSALF 1680
 QY 1661 SIEVQVTIGFGRWATECEPLAILILIVONIVGLMINAIMGCIEMKTAQAHRAETLIF 1720
 DB 1661 SIEVQVTIGFGRWATECEPLAILILIVONIVGLMINAIMGCIEMKTAQAHRAETLIF 1720
 QY 1681 SIEVQVTIGFGRWATECEPLAILILIVONIVGLMINAIMGCIEMKTAQAHRAETLIF 1740
 DB 1681 SIEVQVTIGFGRWATECEPLAILILIVONIVGLMINAIMGCIEMKTAQAHRAETLIF 1740
 QY 1721 SKHAVITLRHGRCLFMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDI PMENG 1780
 DB 1721 SKHAVITLRHGRCLFMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDI PMENG 1780
 QY 1741 SKHAVITLRHGRCLFMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDI PMENG 1800
 DB 1741 SKHAVITLRHGRCLFMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDI PMENG 1800
 QY 1781 VGGNGIFLVAFLIYHYIDSNSPLDYDALPSLHHQDLEIIVILEGVETTGITTOARTS 1840
 DB 1781 VGGNGIFLVAFLIYHYIDSNSPLDYDALPSLHHQDLEIIVILEGVETTGITTOARTS 1840
 QY 1801 VGGNGIFLVAFLIYHYIDSNSPLDYDALPSLHHQDLEIIVILEGVETTGITTOARTS 1860
 DB 1801 VGGNGIFLVAFLIYHYIDSNSPLDYDALPSLHHQDLEIIVILEGVETTGITTOARTS 1860
 QY 1841 YLADEIILMGQFVPIVAEBEDRSVDSKPCNTIKVPTPLCTARQLEDRESLDALTLAS 1900
 DB 1841 YLADEIILMGQFVPIVAEBEDRSVDSKPCNTIKVPTPLCTARQLEDRESLDALTLAS 1900
 QY 1901 SRGPLRAASVAVAKAKPKFSISPDLS 1927
 DB 1901 SRGPLRAASVAVAKAKPKFSISPDLS 1927
 QY 1921 SRGPLRKSVAVAKAKPKFSISPDLS 1947
 DB 1921 SRGPLRKSVAVAKAKPKFSISPDLS 1947

RESULT 4
 US-10-665-283-4
 ; Sequence 4, Application US/10665283
 ; Publication No. US20050063989A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
 ; APPLICANT: DERAND, Renaud
 ; APPLICANT: GARCIA, Elisabeth


```

; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLOUP, Jean
; APPLICANT: VIVAUOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: P63US104
; CURRENT APPLICATION NUMBER: US/10/665, 283
; CURRENT FILING DATE: 2003-09-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1891
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-4

Query Match      98.3%; Score 9734; DB 5; Length 1891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALRGFCADGSDPLDMANVTWNTSNPDFKCFONTLVWVPCFYLMACPFYFLYSRH 60
Db      1 MALRGFCADGSDPLDMANVTWNTSNPDFKCFONTLVWVPCFYLMACPFYFLYSRH 60

Qy      61 DRGYIQMTPLNKTKTALGFLMIVCMADLFYSFWEBSRGIFLAVFLVSPTLIGITTLA 120
Db      61 DRGYIQMTPLNKTKTALGFLMIVCMADLFYSFWEBSRGIFLAVFLVSPTLIGITTLA 120

Qy      121 TPLIQLERRKGVSSGIMLTFWLVALCALALIRSKMTALKEDAQVDLFRDITFYVFS 180
Db      121 TPLIQLERRKGVSSGIMLTFWLVALCALALIRSKMTALKEDAQVDLFRDITFYVFS 180

Qy      181 LLLIQLVLSGCFSDRSPLFSETIHDNCPRESSASFISRTFPMWTGLIVGYQPLEGSD 240
Db      181 LLLIQLVLSGCFSDRSPLFSETIHDNCPRESSASFISRTFPMWTGLIVGYQPLEGSD 240

Qy      241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVVYVSSKDPAPQKSSKVDANEVEAL 300
Db      241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVVYVSSKDPAPQKSSKVDANEVEAL 300

Qy      301 IVSPKQKMPNPLFKVLVYKTFPGPYFLMSFFPKAIDHLMFSGPOILKLIKPVNDYAPD 360
Db      301 IVSPKQKMPNPLFKVLVYKTFPGPYFLMSFFPKAIDHLMFSGPOILKLIKPVNDYAPD 360

Qy      361 WQGYFTVVLVFTVACIOTLVLHOVFIHICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420
Db      361 WQGYFTVVLVFTVACIOTLVLHOVFIHICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420

Qy      421 GEIVNLMSVDAORFMDLATYINMIWSAPLQVILALYLLMLNTGFSVLAVGVAVWVLMPVN 480
Db      421 GEIVNLMSVDAORFMDLATYINMIWSAPLQVILALYLLMLNTGFSVLAVGVAVWVLMPVN 480

Qy      481 AVAMAKTKTYQVAHMSKDNRIKLMNEILINGIKVLKLYAMELAFKDKVLAIRQBELKVLK 540
Db      481 AVAMAKTKTYQVAHMSKDNRIKLMNEILINGIKVLKLYAMELAFKDKVLAIRQBELKVLK 540

Qy      541 KSAVLSAVGFTVNCPTPELVALCTPAVYVITIDENNLIDATAPVSLAEHILAFPLNLP 600
Db      541 KSAVLSAVGFTVNCPTPELVALCTPAVYVITIDENNLIDATAPVSLAEHILAFPLNLP 600

Qy      601 MVAISIVQASVSLKRLIFLSHEELBPSIERRPVKQGGGNSITVYNAFTVARSDEPT 660
Db      601 MVAISIVQASVSLKRLIFLSHEELBPSIERRPVKQGGGNSITVYNAFTVARSDEPT 660

Qy      661 LNSITTSIPGALVAVVGVGVCGKSSLSALLAEMDVVEGVVAIKGSVAVVPOAMVQND 720
Db      661 LNSITTSIPGALVAVVGVGVCGKSSLSALLAEMDVVEGVVAIKGSVAVVPOAMVQND 720

Qy      721 STRENTLIFGQLEPPYRSVIOACALPDEILPSGRTIEIGKGVNLGGGQORVSLAR 780
Db      721 STRENTLIFGQLEPPYRSVIOACALPDEILPSGRTIEIGKGVNLGGGQORVSLAR 780

Qy      781 AVYSNADIVLFDPLSLAVDAVGHKIFENYIGPKMKLNKTRLLVTHSMYSLPQVDVYIV 840
Db      781 AVYSNADIVLFDPLSLAVDAVGHKIFENYIGPKMKLNKTRLLVTHSMYSLPQVDVYIV 840
```

```

Db      781 AVYSNADIVLFDPLSLAVDAVGHKIFENYIGPKMKLNKTRLLVTHSMYSLPQVDVYIV 840
Qy      841 MSGGKISEMSXOEILARDAPAEPLRTVYSTBOEOAEENGVTGSGPKKAKOMNGM 900
Db      841 MSGGKISEMSXOEILARDAPAEPLRTVYSTBOEOAEENGVTGSGPKKAKOMNGM 900

Qy      901 LVTDSAGKQORQLSSSSSYSGDISRHHNSTAEIQKAEKKEFTWKLMKADKQOTGVKL 960
Db      901 LVTDSAGKQORQLSSSSSYSGDISRHHNSTAEIQKAEKKEFTWKLMKADKQOTGVKL 960

Qy      961 SVYWDYWKALGLEIFSLISLFMCNHYASALASNYMSLWTDPIVNGTOEHTVRLSVYG 1020
Db      961 SVYWDYWKALGLEIFSLISLFMCNHYASALASNYMSLWTDPIVNGTOEHTVRLSVYG 1020

Qy      1021 ALGISQIAVFGYSMAVSTGIIASRCLHNDLHSLIRSMSPFERPSSGNLVNREKEL 1080
Db      1021 ALGISQIAVFGYSMAVSTGIIASRCLHNDLHSLIRSMSPFERPSSGNLVNREKEL 1080

Qy      1081 DTVDMSIPEVYIKMFMSLFWVIGACIVILATPIAIIIPPLGIYFPVORFYVASSROL 1140
Db      1081 DTVDMSIPEVYIKMFMSLFWVIGACIVILATPIAIIIPPLGIYFPVORFYVASSROL 1140

Qy      1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDIKVDENOKAYVPSIVANRWLA 1200
Db      1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDIKVDENOKAYVPSIVANRWLA 1200

Qy      1201 VRLCEGNCIVLFAALFAVYSRHSLSAGLVGSLVSLSQVTTTLNMLVRRSMSEMETNIVA 1260
Db      1201 VRLCEGNCIVLFAALFAVYSRHSLSAGLVGSLVSLSQVTTTLNMLVRRSMSEMETNIVA 1260

Qy      1261 VERLKEYSETKEAPMOIOETAPSSMPQVGRVFPNRYCLARYEDLDQFVLRHINVTNGG 1320
Db      1261 VERLKEYSETKEAPMOIOETAPSSMPQVGRVFPNRYCLARYEDLDQFVLRHINVTNGG 1320

Qy      1321 EKVIGVGRTAGKGSLLTGLFRINESABEGIIIDGINIAKIGLHDLRFKTIIPQEPVLF 1380
Db      1321 EKVIGVGRTAGKGSLLTGLFRINESABEGIIIDGINIAKIGLHDLRFKTIIPQEPVLF 1380

Qy      1381 SCSLRNMLDPPSOYSDBEWTSLLELAHKDFVSLPDKLDHCECAEGGENISVQORQLVCL 1440
Db      1381 SCSLRNMLDPPSOYSDBEWTSLLELAHKDFVSLPDKLDHCECAEGGENISVQORQLVCL 1440

Qy      1441 ABALLKRTKLVLDENFAVNDLETDLIOSTIRTOEDCTVLTIARLNTIMYTRIVL 1500
Db      1441 ABALLKRTKLVLDENFAVNDLETDLIOSTIRTOEDCTVLTIARLNTIMYTRIVL 1500

Qy      1501 DKGEIOEGAPSDLLQORGLFYSMAKQAGLVGGGGGMLSRKGIIPBEYVLTBLAEDPAE 1560
Db      1501 DKGEIOEGAPSDLLQORGLFYSMAKQAGLVGGGGGMLSRKGIIPBEYVLTBLAEDPAE 1560

Qy      1561 PRYTRERBARFVSKKGNCAVAKNIREQGRFLQDVFTLLVDLKWPHITLLIFMSPLCSW 1620
Db      1561 PRYTRERBARFVSKKGNCAVAKNIREQGRFLQDVFTLLVDLKWPHITLLIFMSPLCSW 1620

Qy      1621 LLFAMVWMLIAFAGHGLAGEGTVPCVTSIHSSSAPLFSLEVQYTIIGGSGMVEBEP 1680
Db      1621 LLFAMVWMLIAFAGHGLAGEGTVPCVTSIHSSSAPLFSLEVQYTIIGGSGMVEBEP 1680

Qy      1681 LAILILIVONIVGIMINAIMLGCIFMKTAQAHRAETLLFSKAVVTLTBHGRCLCFMLRVG 1740
Db      1681 LAILILIVONIVGIMINAIMLGCIFMKTAQAHRAETLLFSKAVVTLTBHGRCLCFMLRVG 1740

Qy      1741 DLKXSMISATTHMOVVRKTTSPBEGVVPPLHOVDIPMENGVGNGIFLVAAPLIYVVIDS 1800
Db      1741 DLKXSMISATTHMOVVRKTTSPBEGVVPPLHOVDIPMENGVGNGIFLVAAPLIYVVIDS 1800

Qy      1801 NSPLVDLAPSDLHHQDLRIYVLBEGVETGTTTQAFYSYLADELILMGORFPVPIAEBD 1860
Db      1801 NSPLVDLAPSDLHHQDLRIYVLBEGVETGTTTQAFYSYLADELILMGORFPVPIAEBD 1860

Qy      1861 GRYSVDYSKRGNTIKVPTPLCTARQLEBDRS 1891
Db      1861 GRYSVDYSKRGNTIKVPTPLCTARQLEBDRS 1891
```

RESULT 5
US-09-939-853A-86
; Sequence 86, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Buzgeas et al.
; TITLE OF INVENTION: No. US20040039163A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-86

Query Match 79.4%; Score 7860; DB 3; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCASDSDPLMDMNVMTNTSNPDTKCFQNTVLVWPCFYLMACFPFFLYLSRH 60
DB 1 MALRGFCASDSDPLMDMNVMTNTSNPDTKCFQNTVLVWPCFYLMACFPFFLYLSRH 60
QY 61 DRGIQWTPLNKKTALGFLIMVCMADLFYSFWERSRGITLAPVFLVSPLLGITTLLA 120
DB 61 DRGIQWTPLNKKTALGFLIMVCMADLFYSFWERSRGITLAPVFLVSPLLGITTLLA 120
QY 121 TFLQLRRKRGVSSGIMLTFMVALVCAALILRSKMTALKEBAQVDLFPDITFFVYFS 180
DB 121 TFLQLRRKRGVSSGIMLTFMVALVCAALILRSKMTALKEBAQVDLFPDITFFVYFS 180
QY 181 LLLIQLVLSGFSDDSPLESETIHDNPPCESSASFLSRITFMWITGLIVRGYROPLEGSD 240
DB 181 LLLIQLVLSGFSDDSPLESETIHDNPPCESSASFLSRITFMWITGLIVRGYROPLEGSD 240
QY 241 LWSLINKEDTSEQVVPVLVKNKKCECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
DB 241 LWSLINKEDTSEQVVPVLVKNKKCECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPQKEMWNSLFRKLVYKTFGPFYLMSPFKAIHDLMMFSGPOLKLLIFVNDTAPD 360
DB 301 IVKSPQKEMWNSLFRKLVYKTFGPFYLMSPFKAIHDLMMFSGPOLKLLIFVNDTAPD 360
QY 361 WQGFYFVTLFVTAQCLQTLVLAHQFHLICFVSGMRITKTAIVAGVYRKALVITNSARKSSTV 420
DB 361 WQGFYFVTLFVTAQCLQTLVLAHQFHLICFVSGMRITKTAIVAGVYRKALVITNSARKSSTV 420
QY 421 GEIVNIMSVDARFMDLATYINMIWSAPLQVILALYLLMLNIGPSVLGAVAVMLVMPVN 480
DB 421 GEIVNIMSVDARFMDLATYINMIWSAPLQVILALYLLMLNIGPSVLGAVAVMLVMPVN 480
QY 481 AVAMMKTKTYOVAMKSKDNRIKLMNELINGIKYLKLYAMBLAFPDKYLALROBELKYLK 540
DB 481 AVAMMKTKTYOVAMKSKDNRIKLMNELINGIKYLKLYAMBLAFPDKYLALROBELKYLK 540
QY 541 KSAIYLSAVGTFTWCTPFLVALCTFAVVYTIIDENNIIDAQTAFYSLALENTLRPLNITP 600
DB 541 KSAIYLSAVGTFTWCTPFLVALCTFAVVYTIIDENNIIDAQTAFYSLALENTLRPLNITP 600
QY 601 MVISSIVQASVSLRLRIFLSHELEPDSTERRPVKGGTNSITVRNATFTWARSDEPT 660

DB 601 MVISSIVQASVSLRLRIFLSHELEPDSTERRPVKGGTNSITVRNATFTWARSDEPT 660
QY 661 INGITSIPREGALVAVVGVCCKSSLLSALLAEMDVEEAGVAIKGSVAIVYPOQAWIOND 720
DB 661 INGITSIPREGALVAVVGVCCKSSLLSALLAEMDVEEAGVAIKGSVAIVYPOQAWIOND 720
QY 721 SLRENTLIFGQLEBPYRSVIOACALLPDLLEILPSGDRTEIGEXGNVLSGGQKORVSLAR 780
DB 721 SLRENTLIFGQLEBPYRSVIOACALLPDLLEILPSGDRTEIGEXGNVLSGGQKORVSLAR 780
QY 781 AVYSNADIYLPDDPLSANDAHVGHIFENYIGRKGMKKTTRILLVTHSMSTYLPQVDVIY 840
DB 781 AVYSNADIYLPDDPLSANDAHVGHIFENYIGRKGMKKTTRILLVTHSMSTYLPQVDVIY 840
QY 841 MSGGKISEMSYQELARDAFAEFLRTYASTBOEBAENGVTVSGPGKEAKOMENGM 900
DB 841 MSGGKISEMSYQELARDAFAEFLRTYASTBOEBAENGVTVSGPGKEAKOMENGM 900
QY 901 LVTDAGKOLQROLSSSSYSGDISRRHNSYAELOKAEAKKETWKLMEADKAQTGVKL 960
DB 901 LVTDAGKOLQROLSSSSYSGDISRRHNSYAELOKAEAKKETWKLMEADKAQTGVKL 960
QY 961 SYWMDYKKAIGLFISPLSIFLFCMCHVSAASNYSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SYWMDYKKAIGLFISPLSIFLFCMCHVSAASNYSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 961 SYWMDYKKAIGLFISPLSIFLFCMCHVSAASNYSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SYWMDYKKAIGLFISPLSIFLFCMCHVSAASNYSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQIAVFGYSMAVSIIGIILASRCIAVDLHSLRSPSPFERPPSGNLVNRFSKEL 1080
DB 1021 ALGISQIAVFGYSMAVSIIGIILASRCIAVDLHSLRSPSPFERPPSGNLVNRFSKEL 1080
QY 1081 DTVDMSIPEVIKFMGSLFNVIGACIVLLATPIAIIIPPLGIYFPVQRFYVASSROL 1140
DB 1081 DTVDMSIPEVIKFMGSLFNVIGACIVLLATPIAIIIPPLGIYFPVQRFYVASSROL 1140
QY 1141 KRLSEVSRSRYSHFNTELLGVSYTRAFEEQERFIHOSDLKVDENQAAVYPSIYANRWLA 1200
DB 1141 KRLSEVSRSRYSHFNTELLGVSYTRAFEEQERFIHOSDLKVDENQAAVYPSIYANRWLA 1200
QY 1201 VRLCEVNCIYLPALFAVISRHSLSAGLVGSYSLOVTTYINMLVRRSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALFAVISRHSLSAGLVGSYSLOVTTYINMLVRRSSEMETNIVA 1260
QY 1261 VERLKEVSETEKEAPWQIOETAPSSWPQYGRVEFRNYCLRYEDDLFVLRIHNVITNGG 1320
DB 1261 VERLKEVSETEKEAPWQIOETAPSSWPQYGRVEFRNYCLRYEDDLFVLRIHNVITNGG 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINESABGEIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSSLTGLFRINESABGEIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
QY 1381 SGLSRMLDPPSOYSDDEWVTSLELAHLKDFVSALPDKLHECAGEGENTSVGQROLVCL 1440
DB 1381 SGLSRMLDPPSOYSDDEWVTSLELAHLKDFVSALPDKLHECAGEGENTSVGQROLVCL 1440
QY 1441 ABALRLRTKTLVLDENATAVVDLETDLLIOSTIRTOPEDCVTLTAHRLNTIMDTTRVIVL 1500
DB 1441 ABALRLRTKTLVLDENATAVVDLETDLLIOSTIRTOPEDCVTLTAHRLNTIMDTTRVIVL 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDILQORGLFYSMAKDAGLV 1531

RESULT 6
US-10-667-891-6
; Sequence 6, Application US/10667891
; Publication No. US20040171024A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BREY, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRAILLES, MARINE

APPLICANT: RZHETSKY, ANDREY
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
FILE OF INVENTION: ANOPHELES
FILE REFERENCE: 03495-0294-0000
CURRENT APPLICATION NUMBER: US/10/667, 891
PRIORITY FILING DATE: 2003-09-23
PRIORITY APPLICATION NUMBER: 60/413, 469
PRIORITY FILING DATE: 2002-09-26
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO: 6
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo sapiens
US-10-667-891-6

Query Match 79.4%; Score 7860; DB 4; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRFGSADSDMLDMNTMTNTPDFTKCFONTLVWPCYLMACFPFYLYSRH 60
DB 1 MALRFGSADSDMLDMNTMTNTPDFTKCFONTLVWPCYLMACFPFYLYSRH 60
QY 61 DRGYIOMTPLNKTALGFLMIYCNADLFYSFWEKSGIFLAVPLVSPFLIGITLLA 120
DB 61 DRGYIOMTPLNKTALGFLMIYCNADLFYSFWEKSGIFLAVPLVSPFLIGITLLA 120
QY 121 TFLIOLERRKGVSGGIMLTFMVALVICALIILRSKIMTALKEADQVDFPDITFYVFS 180
DB 121 TFLIOLERRKGVSGGIMLTFMVALVICALIILRSKIMTALKEADQVDFPDITFYVFS 180
QY 181 LLLIQLVLSGSDSPFSETIHDNPNCPRESSASFLSRTITWTGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLSGSDSPFSETIHDNPNCPRESSASFLSRTITWTGLIVRGYRPLEGSD 240
QY 181 LLLIQLVLSGSDSPFSETIHDNPNCPRESSASFLSRTITWTGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLSGSDSPFSETIHDNPNCPRESSASFLSRTITWTGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSEQVAVLVKMKKSCAKTRKOPVAVVSSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVAVLVKMKKSCAKTRKOPVAVVSSKDPAPKESKVDANEVEAL 300
QY 301 IVKSPQKEMNPSPLEKVLKTFGPFYFLMSFFPKAIDHLMFSGPOLILKILKFNVDTRAPD 360
DB 301 IVKSPQKEMNPSPLEKVLKTFGPFYFLMSFFPKAIDHLMFSGPOLILKILKFNVDTRAPD 360
QY 361 WOGFFYVVLVPTACLOTVLHQYFHICFVSGMRKTAIVGAYRKALVITNSARKSSTV 420
DB 361 WOGFFYVVLVPTACLOTVLHQYFHICFVSGMRKTAIVGAYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAOREMDLATYINMWSAPLOVITLALYLLMLNGSVLAGVAVMVLAMPVN 480
DB 421 GEIYNLMSVDAOREMDLATYINMWSAPLOVITLALYLLMLNGSVLAGVAVMVLAMPVN 480
QY 481 AVNAMKTKTYQVAMKSKDNRIKLMNEILNGIKVLKYLAMELAFKDKVLAIRBELKVLK 540
DB 481 AVNAMKTKTYQVAMKSKDNRIKLMNEILNGIKVLKYLAMELAFKDKVLAIRBELKVLK 540
QY 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITDENNIIDAOYFVSLFENILRPPLNLP 600
DB 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITDENNIIDAOYFVSLFENILRPPLNLP 600
QY 601 MVISSIVQASVSLKRLIFLSHELEPDSIERPVKCGGCTNSTVNNAFTFARSPPT 660
DB 601 MVISSIVQASVSLKRLIFLSHELEPDSIERPVKCGGCTNSTVNNAFTFARSPPT 660
QY 661 LINGITFISIPGALVAVVGVGCGKSLSLALAMEKVEGVAIKGSVAVVPOQAWIOND 720
DB 661 LINGITFISIPGALVAVVGVGCGKSLSLALAMEKVEGVAIKGSVAVVPOQAWIOND 720
QY 721 SLRENIIFGQLEPPYRSYIACALLPDLIELPSGRTIEGEGVNLSSGQKORVSLAR 780
DB 721 SLRENIIFGQLEPPYRSYIACALLPDLIELPSGRTIEGEGVNLSSGQKORVSLAR 780
QY 781 AVYSNADIYFDPLSLAVDAHVGHIFENYIGPKMKLKNTRILVTHSMYSYLPQVDYIIV 840

DB 781 AVYSNADIYFDPLSLAVDAHVGHIFENYIGPKMKLKNTRILVTHSMYSYLPQVDYIIV 840
QY 841 MSGGKISEMGSYOELARDGAFELFTYASTEOEADAENGVTGSGPKGAOKOMENG 900
DB 841 MSGGKISEMGSYOELARDGAFELFTYASTEOEADAENGVTGSGPKGAOKOMENG 900
QY 901 LYNDSACKOIOROLSSSSSGDLSRHNSIAELQKAEKEETWKIEMADKAOYGVL 960
DB 901 LYNDSACKOIOROLSSSSSGDLSRHNSIAELQKAEKEETWKIEMADKAOYGVL 960
QY 961 SVYWDYKAIQGLFISIFLFCMCHVASALASNYMLSLMTDDPIVNGTOEHTKRLSYG 1020
DB 961 SVYWDYKAIQGLFISIFLFCMCHVASALASNYMLSLMTDDPIVNGTOEHTKRLSYG 1020
QY 1021 ALGISQGIAYFGYMAVISIGIILASRCLHVDLHSLRSFMSFFERTPSGNLVNFSKEL 1080
DB 1021 ALGISQGIAYFGYMAVISIGIILASRCLHVDLHSLRSFMSFFERTPSGNLVNFSKEL 1080
QY 1081 DTVDMSIPEVITKMGSLFNVIACIYIILATPIAIIIPPLGIYFFVQRPYVASSROL 1140
DB 1081 DTVDMSIPEVITKMGSLFNVIACIYIILATPIAIIIPPLGIYFFVQRPYVASSROL 1140
QY 1141 KRLESVSRSPVSHFNFTLGVSVYIRAFEOERFIHOSDLKVDENQKAYPSIVANRMLA 1200
DB 1141 KRLESVSRSPVSHFNFTLGVSVYIRAFEOERFIHOSDLKVDENQKAYPSIVANRMLA 1200
QY 1201 VRLCVCNCIVLPALFAVIRSRHLSAGLVGLSVSLQVTTYLMNLVRRSSEMETNIVA 1260
DB 1201 VRLCVCNCIVLPALFAVIRSRHLSAGLVGLSVSLQVTTYLMNLVRRSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFNRYCLRYREDDLPVLRHINVTING 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFNRYCLRYREDDLPVLRHINVTING 1320
QY 1321 EKVGIWRTAGKSLTLGLFRINESABEGIIIDGINIAKIGLHDAEFKTTIIPQDVLV 1380
DB 1321 EKVGIWRTAGKSLTLGLFRINESABEGIIIDGINIAKIGLHDAEFKTTIIPQDVLV 1380
QY 1381 SGLSRNMLDPPSOYSDSEEVTSLELAHLKDFVSLPDKLHECAGEGENTSVGQROLVCL 1440
DB 1381 SGLSRNMLDPPSOYSDSEEVTSLELAHLKDFVSLPDKLHECAGEGENTSVGQROLVCL 1440
QY 1441 ARALLKRTKLVLDERTAAVDLETDDLIOSTIRTOFEDCTVLIARHLANTIMDYRIYL 1500
DB 1441 ARALLKRTKLVLDERTAAVDLETDDLIOSTIRTOFEDCTVLIARHLANTIMDYRIYL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMADAGLV 1531
DB 1501 DKGEIOEGAPSDLLQORGLFYSMADAGLV 1531

RESULT 7
US-10-607-466-6
Sequence 6, Application US/10807466
Publication No. US20040244066A1
GENERAL INFORMATION:
APPLICANT: ROTH, CHARLES W.
APPLICANT: BREV, PAUL T.
APPLICANT: HOLM, INGE
APPLICANT: GRALLIES, MARINE
APPLICANT: RZHETSKY, ANDREY
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
FILE REFERENCE: 03495-0294-01000
CURRENT APPLICATION NUMBER: US/10/807, 466
CURRENT FILING DATE: 2004-03-24
PRIORITY APPLICATION NUMBER: 10/667, 891
PRIORITY FILING DATE: 2003-09-23
PRIORITY APPLICATION NUMBER: 60/413, 469
PRIORITY FILING DATE: 2002-09-26
NUMBER OF SEQ ID NOS: 139
SOFTWARE: PatentIn Ver. 3.2

```
; SEQ ID NO 6
; LENGTH: 1531
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-807-466-6

Query Match      79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 MALRGFCAGDSDPLMDMNNVTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYLSRH 60
Db      1 MALRGFCAGDSDPLMDMNNVTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYLSRH 60
Oy      61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGTTTLA 120
Db      61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGTTTLA 120
Oy      121 TFLIOLERRKGVOSGIMLTFWVALVCAALILRSKIMTALKEDAOVDLFRDITFYVFS 180
Db      121 TFLIOLERRKGVOSGIMLTFWVALVCAALILRSKIMTALKEDAOVDLFRDITFYVFS 180
Oy      181 LLLIQLVUSCFSDRSPLFSETHIDNPCESSASFLSITTFWITGLIVRGYROPLEGGSD 240
Db      181 LLLIQLVUSCFSDRSPLFSETHIDNPCESSASFLSITTFWITGLIVRGYROPLEGGSD 240
Oy      241 LMSINKEDTSQOVVPLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKYDANEVEAL 300
Db      241 LMSINKEDTSQOVVPLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKYDANEVEAL 300
Oy      301 IVKSPQKMNPSLFPVLYKTGPYFLMSPFKAIHDLMMFSGPOLIKLIFVNDTKAPD 360
Db      301 IVKSPQKMNPSLFPVLYKTGPYFLMSPFKAIHDLMMFSGPOLIKLIFVNDTKAPD 360
Oy      361 MOGYFYTVLFTYACLOTLYVHOYFHCIFVSGMRIKTAVIGAVYKALVITNSAKKSTV 420
Db      361 MOGYFYTVLFTYACLOTLYVHOYFHCIFVSGMRIKTAVIGAVYKALVITNSAKKSTV 420
Oy      421 GEIVNLSVDAQRFNDLTYINMISAPLOVILALYLLMLNGSPVLGAVAVMLMVEVN 480
Db      421 GEIVNLSVDAQRFNDLTYINMISAPLOVILALYLLMLNGSPVLGAVAVMLMVEVN 480
Oy      481 AVAMAKTKTYOVAAHKKSKONRIKLMNEILINGIKVLKYAMBLAFOKVLAIRQBELKYLK 540
Db      481 AVAMAKTKTYOVAAHKKSKONRIKLMNEILINGIKVLKYAMBLAFOKVLAIRQBELKYLK 540
Oy      541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTIDENNILDAQTAFFSLAFNLIRPFLNLP 600
Db      541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTIDENNILDAQTAFFSLAFNLIRPFLNLP 600
Oy      601 MVISIVQASVSLKRLIFLSHEBELPDSIERRPVKDGGGNTSITVRNATFTWASDPT 660
Db      601 MVISIVQASVSLKRLIFLSHEBELPDSIERRPVKDGGGNTSITVRNATFTWASDPT 660
Oy      661 LINGITFSIPEGALVAVVQVCGCKSSLSALLAENDKYEGHVAITGSVAAYVPOQAMIOND 720
Db      661 LINGITFSIPEGALVAVVQVCGCKSSLSALLAENDKYEGHVAITGSVAAYVPOQAMIOND 720
Oy      721 SLRENIILGCOLEEEYRSVIOACALLPDLLEILPSGDRTELGEKVNISGGCKQVSLAR 780
Db      721 SLRENIILGCOLEEEYRSVIOACALLPDLLEILPSGDRTELGEKVNISGGCKQVSLAR 780
Oy      781 AVYSNADIYLPDDPLSAVDAAHVKHIFENVIGPKMKLNKTRILVTTHSKSYLPQVDVIV 840
Db      781 AVYSNADIYLPDDPLSAVDAAHVKHIFENVIGPKMKLNKTRILVTTHSKSYLPQVDVIV 840
Oy      841 MSGKISMGSGYOELLARDGAFAEFLRYASTEGQDAEENGVTGVSQKGEAKOMENGM 900
Db      841 MSGKISMGSGYOELLARDGAFAEFLRYASTEGQDAEENGVTGVSQKGEAKOMENGM 900
Oy      901 LVTTSAGKOLORQSSSSSSSGDISRRHNSTAELOKAKKKETKMLMEADAKQGVYKL 960
Db      901 LVTTSAGKOLORQSSSSSSSGDISRRHNSTAELOKAKKKETKMLMEADAKQGVYKL 960
```

```
Oy      961 SVYWDYKKAIGLFIISFISIFLMCNHVSALASNYWLSLMTDDPIVNGTOEHTKVLRSYVG 1020
Db      961 SVYWDYKKAIGLFIISFISIFLMCNHVSALASNYWLSLMTDDPIVNGTOEHTKVLRSYVG 1020
Oy      1021 ALGISOGIIVAVGYSMAVISIGIILASRCLHYDLHLSIIRSPMSFFERTPSGULVNRFSKEL 1080
Db      1021 ALGISOGIIVAVGYSMAVISIGIILASRCLHYDLHLSIIRSPMSFFERTPSGULVNRFSKEL 1080
Oy      1081 DTVDSMIEPVIKMGMSLFFNVIGACIVILATPIAAIIIPPLGIFYFFVQRFYVASSRQL 1140
Db      1081 DTVDSMIEPVIKMGMSLFFNVIGACIVILATPIAAIIIPPLGIFYFFVQRFYVASSRQL 1140
Oy      1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
Db      1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
Oy      1201 VRLCEVNGCIVLFAALPAVVISRHSLSAGLVGLSVYSLOVTVYLMVLRSSSEMETNIVA 1260
Db      1201 VRLCEVNGCIVLFAALPAVVISRHSLSAGLVGLSVYSLOVTVYLMVLRSSSEMETNIVA 1260
Oy      1261 VERLEKEYSETEKEAPMOIQETAPPSWPOVGRVFEFRNYCLYRREDLDFVLRIHINTING 1320
Db      1261 VERLEKEYSETEKEAPMOIQETAPPSWPOVGRVFEFRNYCLYRREDLDFVLRIHINTING 1320
Oy      1321 EKVGIVGRTGAGKSLTLGLFRINESABGEIIDGINIAKTGLDLRFKTTIIPQDPVLF 1380
Db      1321 EKVGIVGRTGAGKSLTLGLFRINESABGEIIDGINIAKTGLDLRFKTTIIPQDPVLF 1380
Oy      1381 SGSLRMNLDPSSQVSDERWMTSLHLAKDPVSLPDKLHECAEGENLSVGRQOLVCL 1440
Db      1381 SGSLRMNLDPSSQVSDERWMTSLHLAKDPVSLPDKLHECAEGENLSVGRQOLVCL 1440
Oy      1441 ARALLRKTILVLEATAAVDELFTDDLIOGTIRFOPEDCVLTJTAHRLNTIMDYTRVVL 1500
Db      1441 ARALLRKTILVLEATAAVDELFTDDLIOGTIRFOPEDCVLTJTAHRLNTIMDYTRVVL 1500
Oy      1501 DKGEIOEYGAPSDLLQORGLFYSNAXDAGLV 1531
Db      1501 DKGEIOEYGAPSDLLQORGLFYSNAXDAGLV 1531

RESULT 8
US-10-484-577-678
; Sequence 678, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft
; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGT1A
; FILE REFERENCE: F2285PCT-1
; CURRENT APPLICATION NUMBER: US/10/484, 577
; PRIOR FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: PCT/EP 02/08220
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: EP 01 11 7608.8
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: EP 02011710.7
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 683
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 678
; LENGTH: 1531
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-484-577-678

Query Match      79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 61 DRGIOMTPLNKTKTALGFLMTIWCADLFYSFMRSGIFLAPVFLVSPTLGITTLLA 120
DB 61 DRGIOMTPLNKTKTALGFLMTIWCADLFYSFMRSGIFLAPVFLVSPTLGITTLLA 120
QY 121 TFLIQLEBRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDFRDIFFYYVFS 180
DB 121 TFLIQLEBRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDFRDIFFYYVFS 180
QY 181 LLLIQVLVSCFSDSPFLSFETIHDNPNCPRESSASLSTITWMTTGLIVRGYRQPLBESD 240
DB 181 LLLIQVLVSCFSDSPFLSFETIHDNPNCPRESSASLSTITWMTTGLIVRGYRQPLBESD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKKCECAKTRKQPVKVYVSSKDPAPCKESSKVDABEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKKCECAKTRKQPVKVYVSSKDPAPCKESSKVDABEVEAL 300
QY 301 IVKSPCKEMNPSLFKVLYKTEGPFYFLMSFFKAIHDLMPFGPOLKLLIKFVNDYKAPD 360
DB 301 IVKSPCKEMNPSLFKVLYKTEGPFYFLMSFFKAIHDLMPFGPOLKLLIKFVNDYKAPD 360
QY 361 WQGFYVVLVFTVACLQTLVHOFPHICFVSGMRKIKNAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYVVLVFTVACLQTLVHOFPHICFVSGMRKIKNAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMIMSAPLOVITLALYLLMLGSPVLAGVAVMTLAMPVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMIMSAPLOVITLALYLLMLGSPVLAGVAVMTLAMPVN 480
QY 481 AVMAKTKTYOVAMHKSNDNRKIKLMEILNGIKYLLKLYAMELAFKDKVLAIRQELKYLK 540
DB 481 AVMAKTKTYOVAMHKSNDNRKIKLMEILNGIKYLLKLYAMELAFKDKVLAIRQELKYLK 540
QY 541 KSAVLASVGTFTWCTPFLVALCTPFAVYVITDENNIIDAQTAFLFNLIRPLNLP 600
DB 541 KSAVLASVGTFTWCTPFLVALCTPFAVYVITDENNIIDAQTAFLFNLIRPLNLP 600
QY 601 MVLSIYOASVSLKRLIPLISHELEPDSIERPVKCGGNSITVNAFTTARSPPT 660
DB 601 MVLSIYOASVSLKRLIPLISHELEPDSIERPVKCGGNSITVNAFTTARSPPT 660
QY 661 LMGITFSIPGALVAVVGVCGGKSLLSALLAEMDVKEGHVAIKGSVAVYPOQAMQND 720
DB 661 LMGITFSIPGALVAVVGVCGGKSLLSALLAEMDVKEGHVAIKGSVAVYPOQAMQND 720
QY 721 SLRNNIIFGCOLLEPYRVSVOACALLPDLEILPSCDRTIEGKGVNLSGQKQVSLAR 780
DB 721 SLRNNIIFGCOLLEPYRVSVOACALLPDLEILPSCDRTIEGKGVNLSGQKQVSLAR 780
QY 781 AVVSNADIYLPDDPLSAVDAVGHIFENVIGPKGMLKNTKRIILVTHSMSTLPQVDYIV 840
DB 781 AVVSNADIYLPDDPLSAVDAVGHIFENVIGPKGMLKNTKRIILVTHSMSTLPQVDYIV 840
QY 841 MSGGKISEMGSYOELIARDGAFALFRTYASTEOQDAEENGVTGSGPGKEAKOMENG 900
DB 841 MSGGKISEMGSYOELIARDGAFALFRTYASTEOQDAEENGVTGSGPGKEAKOMENG 900
QY 901 LVTDSAGKQLOROLSSSSSYSGDISRHHNSTAELQKAEAKKEETWKLMEADKAQTGVKL 960
DB 901 LVTDSAGKQLOROLSSSSSYSGDISRHHNSTAELQKAEAKKEETWKLMEADKAQTGVKL 960
QY 961 SYVWDVMAKIGLFISEFLFPMCNHVSALASNTWLSLMDDDPVLNCTOEHTRKLSVYG 1020
DB 961 SYVWDVMAKIGLFISEFLFPMCNHVSALASNTWLSLMDDDPVLNCTOEHTRKLSVYG 1020
QY 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHVDLHSLIRSPMSFFERTPSGULVVRFSKEL 1080
DB 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHVDLHSLIRSPMSFFERTPSGULVVRFSKEL 1080
QY 1081 DTUDSMIPEVYKMMGSLFNVIACIVITLATPILAIITIPPLGLIYFFVQAFYASSROL 1140
DB 1081 DTUDSMIPEVYKMMGSLFNVIACIVITLATPILAIITIPPLGLIYFFVQAFYASSROL 1140

QY 1141 KRLSEVSRPVYSHFNETLIGSVYIRAFEEQEREIHOSDLKVDENOKAYPSIVANMLA 1200
DB 1141 KRLSEVSRPVYSHFNETLIGSVYIRAFEEQEREIHOSDLKVDENOKAYPSIVANMLA 1200
QY 1201 VRLCEVNCIVLFAALFAVIRSRHLSAGVLGSYSLOVTTYANMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIVLFAALFAVIRSRHLSAGVLGSYSLOVTTYANMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFRNYCLARREDDLPVLRHINVTING 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFRNYCLARREDDLPVLRHINVTING 1320
QY 1321 EKVGIQVGTAGKSLTLGFRINESAGEIIDIINAIKIGLHDLFKITIIIPQDPVLF 1380
DB 1321 EKVGIQVGTAGKSLTLGFRINESAGEIIDIINAIKIGLHDLFKITIIIPQDPVLF 1380
QY 1381 SGLIRMLDPPSOYSDEEWTSLBLAKDKFVSALPDKLHCEAEGENLSVGQROLVCL 1440
DB 1381 SGLIRMLDPPSOYSDEEWTSLBLAKDKFVSALPDKLHCEAEGENLSVGQROLVCL 1440
QY 1441 ARALRRTKTLVDEATAVADLETDLIQSTIRQFEDCTVLTARHLNTIMDYTRIVL 1500
DB 1441 ARALRRTKTLVDEATAVADLETDLIQSTIRQFEDCTVLTARHLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGAPSDLOQRGLFYSMAXDAGLV 1531
DB 1501 DKGEIOEYGAPSDLOQRGLFYSMAXDAGLV 1531

RESULT 9
US-10-889-503-19
Sequence 19, Application US/10889503
Publication No. US20050063968A1
GENERAL INFORMATION:
APPLICANT: Fox Chase Cancer Center
APPLICANT: Kruh, Gary D.
APPLICANT: Lee, Kun
APPLICANT: Belinsky, Martin G.
TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
FILE REFERENCE: PCCC 98-02
CURRENT APPLICATION NUMBER: US/10/889,503
PRIOR FILING DATE: 2004-07-12
PRIOR APPLICATION NUMBER: US/09/647,140
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US99/06644
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/079,759
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/095,153
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-889-503-19
Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTYNTSNPDPFKCFQNTVLVWVPCFYLMACPPFYFLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTYNTSNPDPFKCFQNTVLVWVPCFYLMACPPFYFLYSRH 60
QY 61 DRGIOMTPLNKTKTALGFLMTIWCADLFYSFMRSGIFLAPVFLVSPTLGITTLLA 120
DB 61 DRGIOMTPLNKTKTALGFLMTIWCADLFYSFMRSGIFLAPVFLVSPTLGITTLLA 120
QY 121 TFLIQLEBRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDFRDIFFYYVFS 180
DB 121 TFLIQLEBRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDFRDIFFYYVFS 180

```

Db      121 TFLIQLERRKVGSSGIMLTFMLVALVCALAILRSKIMTALKEAÖVDLFBDITFYVFS 180
Qy      181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRTFMWITGLIVRGYROPLEBSD 240
Db      181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRTFMWITGLIVRGYROPLEBSD 240
Qy      241 LMSLINKEDTSÖQVVPVIVKMKKECAKTRKÖPVKVVYSSKDPAPÖKSSKYDANEVEAL 300
Db      241 LMSLINKEDTSÖQVVPVIVKMKKECAKTRKÖPVKVVYSSKDPAPÖKSSKYDANEVEAL 300
Qy      301 IVKSPÖKEMNSLFRKVLKTFGPFYFLMSFFPKAIDHLMFMFGPOLIKLLIFVNDTKAPD 360
Db      301 IVKSPÖKEMNSLFRKVLKTFGPFYFLMSFFPKAIDHLMFMFGPOLIKLLIFVNDTKAPD 360
Qy      361 WÖGIFYTYVLFVTAÖLQTLVILHÖYFHIICFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
Db      361 WÖGIFYTYVLFVTAÖLQTLVILHÖYFHIICFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
Qy      421 GEIVNLSVDAÖRFMDLATYINMIWSAPLÖYILALYLLMLNIGPSVLGAVAVMYLMEVFN 480
Db      421 GEIVNLSVDAÖRFMDLATYINMIWSAPLÖYILALYLLMLNIGPSVLGAVAVMYLMEVFN 480
Qy      481 AVAMAKTKTYOVAMKSKDNRIKLMNELLNGIKYLKYAMELAPÖDKYLAIROEELKYLK 540
Db      481 AVAMAKTKTYOVAMKSKDNRIKLMNELLNGIKYLKYAMELAPÖDKYLAIROEELKYLK 540
Qy      541 KSAIYASAVGTFTWVCTPFLVALCTFAVYVTTIDENNILDAÖTAFVSALFNILRPLNLTLP 600
Db      541 KSAIYASAVGTFTWVCTPFLVALCTFAVYVTTIDENNILDAÖTAFVSALFNILRPLNLTLP 600
Qy      601 MVISSIVÖASVSLKRLRIFLSHELEPDSIERRPVKÖGGGNSITVRNATFTWASDPPT 660
Db      601 MVISSIVÖASVSLKRLRIFLSHELEPDSIERRPVKÖGGGNSITVRNATFTWASDPPT 660
Qy      661 LMGITFSPREGALVAVVÖQVCGKSSLLSALLAENDKXEGVAVIIGSAVYPOÖAMIOND 720
Db      661 LMGITFSPREGALVAVVÖQVCGKSSLLSALLAENDKXEGVAVIIGSAVYPOÖAMIOND 720
Qy      721 SLRENIILFCÖLEEPYRSVIOACALPDLLEILPSGDTIEIGEKVNLSGGÖKÖRVSLAR 780
Db      721 SLRENIILFCÖLEEPYRSVIOACALPDLLEILPSGDTIEIGEKVNLSGGÖKÖRVSLAR 780
Qy      781 AVYSNADIYLPDDPLSAVDAHVGHKIFENVIGPKMLNKKTRILLVTHSMSYLPÖVDVILV 840
Db      781 AVYSNADIYLPDDPLSAVDAHVGHKIFENVIGPKMLNKKTRILLVTHSMSYLPÖVDVILV 840
Qy      841 MSGKISMSGYÖELIARDGAFAEFLRTYASTÖBÖDAENGVTGVSÖPGKEAKÖMENG 900
Db      841 MSGKISMSGYÖELIARDGAFAEFLRTYASTÖBÖDAENGVTGVSÖPGKEAKÖMENG 900
Qy      901 LVTSAGKÖLÖRÖLSSSSSYSGDISRHNSSTAELÖKAKKEETKLMBADKÖTÖGYKL 960
Db      901 LVTSAGKÖLÖRÖLSSSSSYSGDISRHNSSTAELÖKAKKEETKLMBADKÖTÖGYKL 960
Qy      961 SVYNDYMKAGILFISFLSILFLMCMNVGALASNYWLSMTDPIVNGTÖHTKVALSYYG 1020
Db      961 SVYNDYMKAGILFISFLSILFLMCMNVGALASNYWLSMTDPIVNGTÖHTKVALSYYG 1020
Qy      1021 ALGISÖGIYAVRGYMAVSIIGILASRCILHVDLHLSILRSPMSFFERTSGNVLNFRSKEL 1080
Db      1021 ALGISÖGIYAVRGYMAVSIIGILASRCILHVDLHLSILRSPMSFFERTSGNVLNFRSKEL 1080
Qy      1081 DTVÖSMIEVILKMFNGSLFENVIGACIVILLATPIAIIIPPLGILYFPÖRÖYVASSÖQL 1140
Db      1081 DTVÖSMIEVILKMFNGSLFENVIGACIVILLATPIAIIIPPLGILYFPÖRÖYVASSÖQL 1140
Qy      1141 KRLBSVRSRÖYVSHENETLGVSVIRAREÖERFIHÖSDLKVDENÖKÖYYPÖSIVANRMLA 1200
Db      1141 KRLBSVRSRÖYVSHENETLGVSVIRAREÖERFIHÖSDLKVDENÖKÖYYPÖSIVANRMLA 1200
Qy      1201 VRLECVNGCIYVPAALFAVISHSHLSAGLVGSYSLOVTTYINMLVMSSEMETNIVA 1260

```

```

Db      1201 VRLECVNGCIYVPAALFAVISHSHLSAGLVGSYSLOVTTYINMLVMSSEMETNIVA 1260
Qy      1261 VERLKEYSETEKEAPWÖIÖETAPSSWÖYGRVEFRNYCLTRYEDDÖVLRHINVTINGG 1320
Db      1261 VERLKEYSETEKEAPWÖIÖETAPSSWÖYGRVEFRNYCLTRYEDDÖVLRHINVTINGG 1320
Qy      1321 EKVGIVERTGAGKSSLTGLFRINESAGEIINDIGINIAKIGLHDLAFKTIIPÖDPVL 1380
Db      1321 EKVGIVERTGAGKSSLTGLFRINESAGEIINDIGINIAKIGLHDLAFKTIIPÖDPVL 1380
Qy      1381 SÖSLRMLDPPSÖYSDÖEWNTSLAHLKÖFVSALPÖKLDEHÖCAGGENLSVÖGÖLÖVCL 1440
Db      1381 SÖSLRMLDPPSÖYSDÖEWNTSLAHLKÖFVSALPÖKLDEHÖCAGGENLSVÖGÖLÖVCL 1440
Qy      1441 ABALIRKTIILVDEATRAVÖDLETDLIÖGSTRÖRÖFÖDÖCVLTIAHRLNTIMDTRYIVL 1500
Db      1441 ABALIRKTIILVDEATRAVÖDLETDLIÖGSTRÖRÖFÖDÖCVLTIAHRLNTIMDTRYIVL 1500
Qy      1501 DKGÖIÖYGAÖSDLLÖÖRGÖLFYSMAKÖAGLV 1531
Db      1501 DKGÖIÖYGAÖSDLLÖÖRGÖLFYSMAKÖAGLV 1531

RESULT 10
US-10-756-149-5033
; Sequence 5033, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Azitec, Alibet
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: Patent version 3.2
; SEQ ID NO 5033
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-5033

Query Match      79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALRGFSGADSDPLMDMNTVMTNSNDFTKCFÖNTVLYVWPÖCYILMACFPÖFÖLYLSRH 60
Db      1 MALRGFSGADSDPLMDMNTVMTNSNDFTKCFÖNTVLYVWPÖCYILMACFPÖFÖLYLSRH 60
Qy      61 DRGYÖMTPLNKKTALGFLIMYCMADLFYSPÖERSGIFLAIVFLVSPILGITTLLA 120
Db      61 DRGYÖMTPLNKKTALGFLIMYCMADLFYSPÖERSGIFLAIVFLVSPILGITTLLA 120
Qy      121 TFLIQLERRKVGSSGIMLTFMLVALVCALAILRSKIMTALKEAÖVDLFBDITFYVFS 180
Db      121 TFLIQLERRKVGSSGIMLTFMLVALVCALAILRSKIMTALKEAÖVDLFBDITFYVFS 180
Qy      181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRTFMWITGLIVRGYROPLEBSD 240
Db      181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRTFMWITGLIVRGYROPLEBSD 240
Qy      241 LMSLINKEDTSÖQVVPVIVKMKKECAKTRKÖPVKVVYSSKDPAPÖKSSKYDANEVEAL 300
Db      241 LMSLINKEDTSÖQVVPVIVKMKKECAKTRKÖPVKVVYSSKDPAPÖKSSKYDANEVEAL 300
Qy      301 IVKSPÖKEMNSLFRKVLKTFGPFYFLMSFFPKAIDHLMFMFGPOLIKLLIFVNDTKAPD 360
Db      301 IVKSPÖKEMNSLFRKVLKTFGPFYFLMSFFPKAIDHLMFMFGPOLIKLLIFVNDTKAPD 360
Qy      361 WÖGIFYTYVLFVTAÖLQTLVILHÖYFHIICFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420

```


Db 361 WQGYFTVLLFVTAQLVTLVHGFHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAQRFMIDLATYINMTMSAPIOVTLALYLLMLNLSGSLAGVAVMTLAMPVN 480
Db 421 GEIYNLMSVDAQRFMIDLATYINMTMSAPIOVTLALYLLMLNLSGSLAGVAVMTLAMPVN 480
Qy 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDLVLAIRQBELKYLK 540
Db 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDLVLAIRQBELKYLK 540
Qy 541 KSAVLASVGTFTWCTPELVALCTPAVYVITIDENNNILDAQTAFLVSLFNLIRPELNL 600
Db 541 KSAVLASVGTFTWCTPELVALCTPAVYVITIDENNNILDAQTAFLVSLFNLIRPELNL 600
Qy 601 MVTSSIVQASVSLRLKIFLSHELEPDSIRRPVYKCGGCTNSITVNNATFTAKRSBPPT 660
Db 601 MVTSSIVQASVSLRLKIFLSHELEPDSIRRPVYKCGGCTNSITVNNATFTAKRSBPPT 660
Qy 661 LMGITFESIPEGALVAVVGVCGCKSSLSALLAEMDKVGHVAKGSVAVYPOQAMTQND 720
Db 661 LMGITFESIPEGALVAVVGVCGCKSSLSALLAEMDKVGHVAKGSVAVYPOQAMTQND 720
Qy 721 SLRENILFGCQLEPPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSGGQKQVSLAR 780
Db 721 SLRENILFGCQLEPPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSGGQKQVSLAR 780
Qy 781 AVVSNADIIYLFDDLSAVDAVGHGIFENVYIGPKMLKNTKRIIVTMSKTLPOVDYIV 840
Db 781 AVVSNADIIYLFDDLSAVDAVGHGIFENVYIGPKMLKNTKRIIVTMSKTLPOVDYIV 840
Qy 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOBDAEENGVTGSGPGKEAKQOMNGM 900
Db 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOBDAEENGVTGSGPGKEAKQOMNGM 900
Qy 901 LVTDSACKOLOROLSSSSSYSGDISRHHNSTAELQKAKKEETWKLMEADKAQTGOVKL 960
Db 901 LVTDSACKOLOROLSSSSSYSGDISRHHNSTAELQKAKKEETWKLMEADKAQTGOVKL 960
Qy 961 SVYWDYMKATGLFISFLSIFLPMCNHVSALASNTWLSMTDDPVLNCTOEBITKRLSVYG 1020
Db 961 SVYWDYMKATGLFISFLSIFLPMCNHVSALASNTWLSMTDDPVLNCTOEBITKRLSVYG 1020
Qy 1021 ALGISOGIAVFGYSMAVSIIGIILASRCILHVDLHSIIRSPMSPEPRTPSGVLNRFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGIILASRCILHVDLHSIIRSPMSPEPRTPSGVLNRFSEKL 1080
Qy 1081 DTVDMSIPEVYIKMFMGSLFNVIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
Db 1081 DTVDMSIPEVYIKMFMGSLFNVIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
Qy 1141 KRLSVSRSPVYSHFNETLLGVSYTRAPEBOERTIHOSDLKVDENQKAYTPSIYANFWLA 1200
Db 1141 KRLSVSRSPVYSHFNETLLGVSYTRAPEBOERTIHOSDLKVDENQKAYTPSIYANFWLA 1200
Qy 1201 VRLCEVCNCTVLPALPAVLSRHSLSAGLVGSYSLOVTTYINMLVRSSSEMETNIVA 1260
Db 1201 VRLCEVCNCTVLPALPAVLSRHSLSAGLVGSYSLOVTTYINMLVRSSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKEKPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTNGG 1320
Db 1261 VERLKEVSETEKEKPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTNGG 1320
Qy 1321 EKVGIIVRTGAGKSSLTLGLFRINESAGEIIIDIGINIAKIGLHDLRFKTIITIPQDVL 1380
Db 1321 EKVGIIVRTGAGKSSLTLGLFRINESAGEIIIDIGINIAKIGLHDLRFKTIITIPQDVL 1380
Qy 1381 SGLRMLDPPSOYSDSEVWTSLELAHKDFVSLPKLDEHCEKAGEENISVQROKIVCL 1440
Db 1381 SGLRMLDPPSOYSDSEVWTSLELAHKDFVSLPKLDEHCEKAGEENISVQROKIVCL 1440
Qy 1441 ARALLRKTCLIVLDEATAVDLETDLIQSTIRTOFEDCTVLATAHRLANTIMDTYIVL 1500
Db 1441 ARALLRKTCLIVLDEATAVDLETDLIQSTIRTOFEDCTVLATAHRLANTIMDTYIVL 1500

Qy 1501 DKGEIOEGAPSDLLQORGLFYSGMAKDGAV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYSGMAKDGAV 1531

RESULT 11
US-09-939-853A-87
; Sequence 87, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgees et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-87

Query Match 78.5%; Score 7769; DB 3; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 DMVNTNNTSPDPTKCRQNTLVWVPCFYLMACFPFYLSHHDGYIOMTPLNKTCTA 76
Db 1 DMVNTNNTSPDPTKCRQNTLVWVPCFYLMACFPFYLSHHDGYIOMTPLNKTCTA 60

Qy 77 LGFLMLIVCWADLFYSFEMERSRGIPLAPVFLVSPTLIGITTLATFLIQLERRKGVOSSG 136
Db 61 LGFLMLIVCWADLFYSFEMERSRGIPLAPVFLVSPTLIGITTLATFLIQLERRKGVOSSG 120

Qy 137 IMTFMLVAVCALILRSKIMTALKEBDAQVDLFRDITFYVYSLIILQLVLSGFSRSP 196
Db 121 IMTFMLVAVCALILRSKIMTALKEBDAQVDLFRDITFYVYSLIILQLVLSGFSRSP 180

Qy 197 LPSETIHDPPCESSASFLSRIITFMWITGLIVRGYRQPLEGSDLSINKEDTSEQVVPV 256
Db 181 LPSETIHDPPCESSASFLSRIITFMWITGLIVRGYRQPLEGSDLSINKEDTSEQVVPV 240

Qy 257 LVNKMKECKAKTRKQPVKYVYSSKDPAPQKSSKVDANEBEVALIVKSPKEMNPISLFKY 316
Db 241 LVNKMKECKAKTRKQPVKYVYSSKDPAPQKSSKVDANEBEVALIVKSPKEMNPISLFKY 300

Qy 317 LYTFGPFYFLMSFFFAIHDLMNFGSPQIILKLIKTAVNDTKADWQGYFTVLLFVTACL 376
Db 301 LYTFGPFYFLMSFFFAIHDLMNFGSPQIILKLIKTAVNDTKADWQGYFTVLLFVTACL 360

Qy 377 QTLVLAHQYFHI CVSGMRKIKTAIVGAVYRKALVITNSARKSSVVGSI VNLMSVDAQRFMD 436
Db 361 QTLVLAHQYFHI CVSGMRKIKTAIVGAVYRKALVITNSARKSSVVGSI VNLMSVDAQRFMD 420

Qy 437 LATYINMISAPQOVLIALYILMLNLSGSLAGVAVMTLAMPVNAVMAKTKTYQVAHMK 496
Db 421 LATYINMISAPQOVLIALYILMLNLSGSLAGVAVMTLAMPVNAVMAKTKTYQVAHMK 480

Qy 497 SKONRIKLMNEILNGIKVLYKLYAMELAFKDLVLAIRQBELKYLKSAVYSAVGTFTWCT 556
Db 481 SKONRIKLMNEILNGIKVLYKLYAMELAFKDLVLAIRQBELKYLKSAVYSAVGTFTWCT 540

Qy 557 PFLVALCTPAVYVITIDENNNILDAQTAFLVSLALFNLIRFNLILPMYISSIVQASVSLKRL 616

```

|||||
Db      541 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSIALFNILRPPLNILPMVSISSIVOASVSLKRL
Qy      617 RIFLSHELEPDSIERRPVYDGGGNSITVRNATFTMARSPPTLNGITFSIPGALVAV
Db      601 RIFLSHELEPDSIERRPVYDGGGNSITVRNATFTMARSPPTLNGITFSIPGALVAV
Qy      677 VGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRNLIIFGQLEBPY
Db      661 VGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRNLIIFGQLEBPY
Qy      737 YRSVIOACALLPDEILIPSGDRTIEGEGVNLSSGQKORVSLARAVYSNADITYLPDPLS
Db      721 YRSVIOACALLPDEILIPSGDRTIEGEGVNLSSGQKORVSLARAVYSNADITYLPDPLS
Qy      797 AVDAHVGHIFENYIGPKGMLKNKTRILVTHSMSTLPQVDYIYVMSGKISBMSYOBLL
Db      781 AVDAHVGHIFENYIGPKGMLKNKTRILVTHSMSTLPQVDYIYVMSGKISBMSYOBLL
Qy      857 ARDGAFAEFLRTYASTEOBODAEENGVTGVSQPKGEAKQOMENGMVLTDSAGKOLQROLSS
Db      841 ARDGAFAEFLRTYASTEOBODAEENGVTGVSQPKGEAKQOMENGMVLTDSAGKOLQROLSS
Qy      917 SSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAOTGOVKLSVYDYMKAIGLFTSF
Db      901 SSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAOTGOVKLSVYDYMKAIGLFTSF
Qy      977 LSTLFMCHNVSALASNYWLSLMTDDPIVNGIOEHTKRLSVYGAIGLSOGIAYVGYMA
Db      961 LSTLFMCHNVSALASNYWLSLMTDDPIVNGIOEHTKRLSVYGAIGLSOGIAYVGYMA
Qy      1037 VSIGIILASRCLAHVDLHSLIRSPMSFFERTPSGNLVNRFSEKEDTVDMSIPEVIKQEMG
Db      1021 VSIGIILASRCLAHVDLHSLIRSPMSFFERTPSGNLVNRFSEKEDTVDMSIPEVIKQEMG
Qy      1097 SLEFVIGACIYIILATPIAIIIPPLGIYFFVQAFYVASSRQKLSVSRSPPYSHFN
Db      1081 SLEFVIGACIYIILATPIAIIIPPLGIYFFVQAFYVASSRQKLSVSRSPPYSHFN
Qy      1157 ETLGVSIVIRAFEEBERFIHOSDLKVDENOKAYPSIVANRMLAVRLCEVGNCTYLPAAL
Db      1141 ETLGVSIVIRAFEEBERFIHOSDLKVDENOKAYPSIVANRMLAVRLCEVGNCTYLPAAL
Qy      1217 FAVISRHSLSAGLVGLSVSYSLOVTTYLNLVLRMSSEMETNIIVAVERLKEYSETEKEAPW
Db      1201 FAVISRHSLSAGLVGLSVSYSLOVTTYLNLVLRMSSEMETNIIVAVERLKEYSETEKEAPW
Qy      1277 QIOETAPSSWPQYGRVFEFRNYCLRYREDLDFVLPHINVTINGEKVGI VGRTGAGKSSL
Db      1261 QIOETAPSSWPQYGRVFEFRNYCLRYREDLDFVLPHINVTINGEKVGI VGRTGAGKSSL
Qy      1337 TLGIFRINESABGEIIDIGINIAKIGLHDLRFKTIIPODPYLPSGSLRMLDPPSOQSD
Db      1321 TLGIFRINESABGEIIDIGINIAKIGLHDLRFKTIIPODPYLPSGSLRMLDPPSOQSD
Qy      1397 BEVWTSLELAHKDFVSALPKLDHECAGGENLSVGOROLVCLARALLRKTILVLDDEA
Db      1381 BEVWTSLELAHKDFVSALPKLDHECAGGENLSVGOROLVCLARALLRKTILVLDDEA
Qy      1457 TAAVLETFDDLIQSTIRFOEDCTVLTIAHRLNTIMDYTRYI VLDKGEIOEGABSDLLQ
Db      1441 TAAVLETFDDLIQSTIRFOEDCTVLTIAHRLNTIMDYTRYI VLDKGEIOEGABSDLLQ
Qy      1517 ORGLFYSAKADAGLV 1531
Db      1501 ORGLFYSAKADAGLV 1515

```

```

RESULT 12
US-10-618-281-42
; Sequence 42, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-281-42

Query Match      78.4%; Score 7763; DB 5; Length 1515;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      17 DNANTNTSNPDFTKCFQNTVYLVWPCFYIMACPFYFLYSRHRDGYIQMTPLNKTKTA
Db      1 DNANTNTSNPDFTKCFQNTVYLVWPCFYIMACPFYFLYSRHRDGYIQMTPLNKTKTA
Qy      77 LGFLMLIVCANADLFYSWERSRGIFLAPVFLVSPTLGITTILATFLIOLERRGVQSSG
Db      61 LGFLMLIVCANADLFYSWERSRGIFLAPVFLVSPTLGITTILATFLIOLERRGVQSSG
Qy      137 IMLTFMLVALVCALAILRSKIMTALKEADAQVDFRDTTFVYFSLILQVLSCFSRSP
Db      121 IMLTFMLVALVCALAILRSKIMTALKEADAQVDFRDTTFVYFSLILQVLSCFSRSP
Qy      197 LFSETHIDPNCPESSASFLSRITFWWITGLIYNGYQPLEGSLMSLNKEDISEQVVPV
Db      181 LFSETHIDPNCPESSASFLSRITFWWITGLIYNGYQPLEGSLMSLNKEDISEQVVPV
Qy      257 LVKWKKECAKTRQOPKVVYSSKDPQOPKESSVVDNNEVEALIVSPQKEMNPSLFKV
Db      241 LVKWKKECAKTRQOPKVVYSSKDPQOPKESSVVDNNEVEALIVSPQKEMNPSLFKV
Qy      317 LYKTFGYFLMSFFFKAIHDLMPGPOILKLIKFNVDTKAPDMOGFYTVLLFVTAQL
Db      301 LYKTFGYFLMSFFFKAIHDLMPGPOILKLIKFNVDTKAPDMOGFYTVLLFVTAQL
Qy      377 QTLVLAHQYFHCFSVGNRIKTAIVIGAYYRKALVITNSARKSIVGEIYVNLMSVDAORFMD
Db      361 QTLVLAHQYFHCFSVGNRIKTAIVIGAYYRKALVITNSARKSIVGEIYVNLMSVDAORFMD
Qy      437 LATYNNIMWSAPLOVITIALYLYLMTNIGPSYIAGVAVNWVAVNANAMKTKTYQVAHMK
Db      421 LATYNNIMWSAPLOVITIALYLYLMTNIGPSYIAGVAVNWVAVNANAMKTKTYQVAHMK
Qy      497 SKDNRKIKAMEIINGIKVILKLYAMELAFKQVLAIROBELKVLKKSAYLSAVGTFTWCT
Db      481 SKDNRKIKAMEIINGIKVILKLYAMELAFKQVLAIROBELKVLKKSAYLSAVGTFTWCT
Qy      557 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSIALFNILRPPLNILPMVSISSIVOASVSLKRL
Db      541 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSIALFNILRPPLNILPMVSISSIVOASVSLKRL
Qy      617 RIFLSHELEPDSIERRPVYDGGGNSITVRNATFTMARSPPTLNGITFSIPGALVAV
Db      601 RIFLSHELEPDSIERRPVYDGGGNSITVRNATFTMARSPPTLNGITFSIPGALVAV
Qy      677 VGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRNLIIFGQLEBPY
Db      661 VGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRNLIIFGQLEBPY
Qy      737 YRSVIOACALLPDEILIPSGDRTIEGEGVNLSSGQKORVSLARAVYSNADITYLPDPLS

```

```

Db 721 YRSVIOACALLPDLLEIPSGDRTEIGEKVNLGGQXORVSLARAVSNADIVLFDPLS 780
Qy 797 AVDAHVGHIFENYIGPKMKJNKRILVTHSMYLPQVDVITVMSGGKISEMSYOBLL 856
Db 781 AVDAHVGHIFENYIGPKMKJNKRILVTHSMYLPQVDVITVMSGGKISEMSYOBLL 840
Qy 857 ARDAFAEFLRTVASTEOQDAENGVTGVSQPKAKQOMENGLVTDASAKOLOROLSS 916
Db 841 ARDAFAEFLRTVASTEOQDAENGVTGVSQPKAKQOMENGLVTDASAKOLOROLSS 900
Qy 917 SSSYSQDISRHNSHTAELOKAEAKKEETWKLMEADKQGTQVLSVYDWMKAIQGLTFSF 976
Db 901 SSSYSQDISRHNSHTAELOKAEAKKEETWKLMEADKQGTQVLSVYDWMKAIQGLTFSF 960
Qy 977 LSTFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVGALGISQIAAFGYSMA 1036
Db 961 LSTFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVGALGISQIAAFGYSMA 1020
Qy 1037 VSIGGILASRCLHVDLHSLIRSPMSFPERTPSGNLVNRFSKELDTVDSMIPEVIKPMFG 1096
Db 1021 VSIGGILASRCLHVDLHSLIRSPMSFPERTPSGNLVNRFSKELDTVDSMIPEVIKPMFG 1080
Qy 1097 SLENVIGACIVILLATPIAAIIPPLGLIYFVORFYVASSROKRLSESVRSRPSYSHN 1156
Db 1081 SLENVIGACIVILLATPIAAIIPPLGLIYFVORFYVASSROKRLSESVRSRPSYSHN 1140
Qy 1157 ETLIGSVIRAFEEQERFIHOSDKVDENOKAYYPSIVANRWLAVLECVNCTVLPAL 1216
Db 1141 ETLIGSVIRAFEEQERFIHOSDKVDENOKAYYPSIVANRWLAVLECVNCTVLPAL 1200
Qy 1217 FAVISRSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETIVAVEBLKESETEKRAPW 1276
Db 1201 FAVISRSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETIVAVEBLKESETEKRAPW 1260
Qy 1277 QIOETAPSSWPQVGRVEFRNYCLRYREDDLVLRHINVTINGEKYGIVRTGAKSSL 1336
Db 1261 QIOETAPSSWPQVGRVEFRNYCLRYREDDLVLRHINVTINGEKYGIVRTGAKSSL 1320
Qy 1337 TLGIFRINESAEGEIIIDGINIAKIGHDLRFKTIIPDPVYFSGSIRNMLDPFSQYSD 1396
Db 1321 TLGIFRINESAEGEIIIDGINIAKIGHDLRFKTIIPDPVYFSGSIRNMLDPFSQYSD 1380
Qy 1397 EEWTSLELAHLKQFVSALPDKDHECAGEGENSVGOROLVCLARALLRKTLLVYDEA 1456
Db 1381 EEWTSLELAHLKQFVSALPDKDHECAGEGENSVGOROLVCLARALLRKTLLVYDEA 1440
Qy 1457 TAAVDLETDLLIOSTIRTOEDCTVLTIAHRLANTIMYTRVIVLDKGEIOEYGA PSDLQ 1516
Db 1441 TAAVDLETDLLIOSTIRTOEDCTVLTIAHRLANTIMYTRVIVLDKGEIOEYGA PSDLQ 1500
Qy 1517 QRGIFYSMACDAGLV 1531
Db 1501 QRGIFYSMACDAGLV 1515

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1718
; LENGTH: 1388
; TYPE: FRT
; ORGANISM: Homo sapiens
us-10-408-765A-1718

Query Match      71.1%; Score 7037; DB 4; Length 1388;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;

Qy 67 MTPLNKTKTALGFLMTIVCWADLFYSFMEERSGIFLAPVPLVSPTLGITTLATFLIQL 126
Db 1 MTPLNKTKTALGFLMTIVCWADLFYSFMEERSGIFLAPVPLVSPTLGITTLATFLIQL 60
Qy 127 ERKGVOSSGIMLTFWLVALVICALAIIRSKIMTKEDAOVDLFRDITFYVYFSLLIQL 186
Db 61 ERKGVOSSGIMLTFWLVALVICALAIIRSKIMTKEDAOVDLFRDITFYVYFSLLIQL 97
Qy 187 VLSCFSRSLPFSFTIHDPNCPRESSAFSLRTFWITGLIYRGYROPLEGSDMLSK 246
Db 98 -----NCPRESSAFSLRTFWITGLIYRGYROPLEGSDMLSK 138
Qy 247 EDTSEQVVPVLVKNMKKECAKTRKQPKVYSSKDPAPQKSSKVDANEVEALIVKSPQ 306
Db 139 EDTSEQVVPVLVKNMKKECAKTRKQPKVYSSKDPAPQKSSKVDANEVEALIVKSPQ 198
Qy 307 KENNPISLFKVLKTYTGPYFLMSFFPKAIDHLMFSGPQILKLIFVNDTKAPDMQGYFY 366
Db 199 KENNPISLFKVLKTYTGPYFLMSFFPKAIDHLMFSGPQILKLIFVNDTKAPDMQGYFY 258
Qy 367 TYLLFTYACIOTLVLYOYHICFVSGKRTAVIYAVYRQALVYTNARSSTYGEIVNL 426
Db 259 TYLLFTYACIOTLVLYOYHICFVSGKRTAVIYAVYRQALVYTNARSSTYGEIVNL 318
Qy 427 MSYDAORFMDLAVYINWISAPLOVITALLVLMNLNTPSYLAGVAVVWLVVAVNAWAMK 486
Db 319 MSYDAORFMDLAVYINWISAPLOVITALLVLMNLNTPSYLAGVAVVWLVVAVNAWAMK 378
Qy 487 TKTYOVAHMKSKONRIKLANEILNGIKVLKLYAMELAFKQKVALIROEBELKVLKKSAYLS 546
Db 379 TKTYOVAHMKSKONRIKLANEILNGIKVLKLYAMELAFKQKVALIROEBELKVLKKSAYLS 438
Qy 547 AVGTFTVCTPRLVALCTPAVYTTIDENNLDAQTAFAVSLATFNILRFPNLIIPMTGSI 606
Db 439 AVGTFTVCTPRLVALCTPAVYTTIDENNLDAQTAFAVSLATFNILRFPNLIIPMTGSI 498
Qy 607 VQASVSLKRLRIFLSHEELPDSIERRPVQDGGTNSITVRNATFTWARSDDPTLNGITF 666
Db 499 VQASVSLKRLRIFLSHEELPDSIERRPVQDGGTNSITVRNATFTWARSDDPTLNGITF 558
Qy 667 SIPEGALVAVVGQVGGKSSLSALLAEMDKVEGHAIVKSVAVVPOQAMVIONDSURENI 726
Db 559 SIPEGALVAVVGQVGGKSSLSALLAEMDKVEGHAIVKSVAVVPOQAMVIONDSURENI 618
Qy 727 LFGCOLBEPYRSVIOACALLPDLLEIPSGDRTEIGEKVNLGGQXORVSLARAVY SNA 786
Db 619 LFGCOLBEPYRSVIOACALLPDLLEIPSGDRTEIGEKVNLGGQXORVSLARAVY SNA 678
Qy 787 DITLPDPLSAVAHAGKHIFENVIGPKMKJNKRILVTHSMYLPQVDVITVMSGGKI 846
Db 679 DITLPDPLSAVAHAGKHIFENVIGPKMKJNKRILVTHSMYLPQVDVITVMSGGKI 738
Qy 847 SEMGSYOELLARDAFAEFLRTVASTEOQDAENGVTGVSQPKAKQOMENGLVTDASA 906
Db 739 SEMGSYOELLARDAFAEFLRTVASTEOQDAENGVTGVSQPKAKQOMENGLVTDASA 798
Qy 907 GKOLQOLSSSSSYSDISRHNSHTAELOKAEAKKEETWKLMEADKQGTQVLSVYDWMY 966
Db 799 GKOLQOLSSSSSYSDISRHNSHTAELOKAEAKKEETWKLMEADKQGTQVLSVYDWMY 858
Qy 967 MKAIGLFIPLSLIFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVGALGISQ 1026

```

```

RESULT 13
US-10-408-765A-1718
; Sequence 1718, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Pabry, Bojn D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408, 765A
; NUMBER OF SEQ ID NOS: 3077

```

Db 859 MKAIGLFIISLIFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKVRSLVYALGISQ 918
Qy 1027 GIAVFGSMASVIGIGILASRCLAHVDLHLSIRSPMSFFERTPSGNLVNRFSEKELDTYDSM 1086
Db 919 GIAVFGSMASVIGIGILASRCLAHVDLHLSIRSPMSFFERTPSGNLVNRFSEKELDTYDSM 978
Qy 1087 IPEYIKPMGSLFNVIAGACIYIILATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSEV 1146
Db 979 IPEYIKPMGSLFNVIAGACIYIILATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSEV 1038
Qy 1147 SRSVYSHFNETLLGVSVIRAFEBRERFIHOSDKVDENOKAYPSIYANMILVRLSEV 1206
Db 1039 SRSVYSHFNETLLGVSVIRAFEBRERFIHOSDKVDENOKAYPSIYANMILVRLSEV 1098
Qy 1207 GNCIVLPAALFAVISRHSLSAGLVGLSVYSISQVTTYINMLVMSSEMETNIIVAVERLKE 1266
Db 1099 GNCIVLPAALFAVISRHSLSAGLVGLSVYSISQVTTYINMLVMSSEMETNIIVAVERLKE 1158
Qy 1267 YSETEKAPMOIOETAPPSSWPQYGRVFPNRYCLRYEDLDPLRHINVTINGGERKGIIV 1326
Db 1159 YSETEKAPMOIOETAPPSSWPQYGRVFPNRYCLRYEDLDPLRHINVTINGGERKGIIV 1218
Qy 1327 GRTAGKSSLTGLFPRINESAGEIIIDGINIAKIGHDLRPKITIIIPQDVLFSGSLRM 1386
Db 1219 GRTAGKSSLTGLFPRINESAGEIIIDGINIAKIGHDLRPKITIIIPQDVLFSGSLRM 1278
Qy 1387 NLDPFSQYSDDEWVTSLELAHKDFVSALPKLDHECAGEGENTL SVGQROLVCLARALLR 1446
Db 1279 NLDPFSQYSDDEWVTSLELAHKDFVSALPKLDHECAGEGENTL SVGQROLVCLARALLR 1338
Qy 1447 KTKILVDEATPAVDLETDLIQSTIRFOFEDCTVLTIAHRLNTIMDYTR 1496
Db 1339 KTKILVDEATPAVDLETDLIQSTIRFOFEDCTVLTIAHRLNTIMDYTR 1388

RESULT 14
US-09-939-853A-88
; Sequence 88, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OR INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 1528
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-88

Query Match 70.7%; Score 7002.5; DB 3; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

Qy 1 MALNGFCGADSGDPLMDNNTVMTNSNDPFTKCFQNTVLVWPCFLVLAACFPFPIYLSRH 60
Db 1 MALNSFCGADSGDPLMDNNTVMTNSNDPFTKCFQNTVLVWPCFLVLAACFPFPIYLSRH 60
Qy 61 DRGIQWMLNKTKTALGFLIMIVCMADLFYSFMRSGIFLAPFVLSPTLLGTTLLA 120
Db 61 DRGIQWMLNKTKTALGFLIMIVCMADLFYSFMRSGIFLAPFVLSPTLLGTTLLA 120

Qy 121 TFLIQERRRGVSGSGIMLFMWLVALYCALAALRSKMTLKEBDAQVDLDRDITFYFYS 180
Db 121 TFLIQERRRGVSGSGIMLFMWLVALYCALAALRSKMTLKEBDAQVDLDRDITFYFYS 180
Qy 181 LLLIQVLSCFSPRLFSETIHDNPPCESASFLSRIFPMWITGLIVGYRQPLEGSD 240
Db 181 LVLQVLSCFSPRLFSETIHDNPPCESASFLSRIFPMWITGLIVGYRQPLEGSD 240
Qy 241 LMSLNKEDTSEOVFVLVKNMKCECATRKQOPVKNYS -SKDPAQPKSSKVDANEEVEA 299
Db 241 LMSLNKEDTSEOVFVLVKNMKCECATRKQOPVKNYS -SKDPAQPKSSKVDANEEVEA 300
Qy 300 LIYKSPQKEMNPISLFLKLYTTFGPFYLMSPFKAIDHLMPSGQIILKLIKPNYDKAP 359
Db 301 LIYKSPQKEMNPISLFLKLYTTFGPFYLMSPFKAIDHLMPSGQIILKLIKPNYDKAP 360
Qy 360 DMQGFYTVLLFVYACIQTLAHQYFHCIVSGSRITAYIGAVYRKALYTNARSST 419
Db 361 DMQGFYTVLLFVYACIQTLAHQYFHCIVSGSRITAYIGAVYRKALYTNARSST 420
Qy 420 VGEIVNMSYDAQRFMDLATYINMWSAPLOVILATYILMLNGPSYLAGAVVWLVNVPV 479
Db 421 VGEIVNMSYDAQRFMDLATYINMWSAPLOVILATYILMLNGPSYLAGAVVWLVNVPV 480
Qy 480 NAYVAMKTKTYQVAHMSKONRIKLANEINANGIKYKLYAMELAFKONVLAIRQEEIKVL 539
Db 481 NAYVAMKTKTYQVAHMSKONRIKLANEINANGIKYKLYAMELAFKONVLAIRQEEIKVL 540
Qy 540 KKSAYLSAVGFPMWCPLVALCTFAVYVTIDENNLIDQOTAFVSLAFILRFPNLIL 599
Db 541 KKSAYLSAVGFPMWCPLVALCTFAVYVTIDENNLIDQOTAFVSLAFILRFPNLIL 600
Qy 600 PMVSSIVQASVSLKRLRIFLSHELEPDSIERPVDGGGWSITVRNATFTWASDDP 659
Db 601 PMVSSIVQASVSLKRLRIFLSHELEPDSIERPVDGGGWSITVRNATFTWASDDP 659
Qy 660 TLNGITFSIPBGALVAVGVGCGKSSLSALLAEMDKRBGHVAKSVAVYVPOQAMION 719
Db 660 TLNGITFSIPBGALVAVGVGCGKSSLSALLAEMDKRBGHVAKSVAVYVPOQAMION 719
Qy 720 DSLRENTLFGQLEEPYRYSYQCALPLELTPSGDRIEIGKGNLSGGQKORVSLA 779
Db 720 DSLRENTLFGQLEEPYRYSYQCALPLELTPSGDRIEIGKGNLSGGQKORVSLA 779
Qy 780 RAYVSNADYILFDDPLSAVDAAHVGKHI FENVYGPCKMLKXKTRILTVHSHMSYILPQVDYII 839
Db 780 RAYVSNADYILFDDPLSAVDAAHVGKHI FENVYGPCKMLKXKTRILTVHSHMSYILPQVDYII 839
Qy 840 VMSGKISSENGSYQELIARDGAFAEFLRTYASTFEQEDAEENGVTGVSGPKKAKOMENG 899
Db 840 VMSGKISSENGSYQELIARDGAFAEFLRTYASTFEQEDAEENGVTGVSGPKKAKOMENG 896
Qy 900 MLVYDASGKQOLRQLSSSSSYSGDISRHHNSTALQKAEKKEETWLMEDKQOTQYK 959
Db 900 MLVYDASGKQOLRQLSSSSSYSGDISRHHNSTALQKAEKKEETWLMEDKQOTQYK 955
Qy 959 MLVYDASGKQOLRQLSSSSSYSGDISRHHNSTALQKAEKKEETWLMEDKQOTQYK 955
Db 959 MLVYDASGKQOLRQLSSSSSYSGDISRHHNSTALQKAEKKEETWLMEDKQOTQYK 955
Qy 960 LSVYMDYKAIIGLFIISLIFLPMCNHVSALASNYWLSLMTDD -PIYNGTOEHTKVRSLV 1018
Db 960 LSVYMDYKAIIGLFIISLIFLPMCNHVSALASNYWLSLMTDD -PIYNGTOEHTKVRSLV 1015
Qy 1016 YGALGILQGAALIFQYMAVSGIGIFAFARRLHLDLVNVLASPMSPFFERTPSGNLVNRFSSK 1075
Db 1016 YGALGILQGAALIFQYMAVSGIGIFAFARRLHLDLVNVLASPMSPFFERTPSGNLVNRFSSK 1075
Qy 1079 ELDTVDNSMIPVYIMFGGSLPNVIGACIYIILATPIAAIIIPPLGLIYFFVQRFYVASSR 1138
Db 1076 ELDTVDNSMIPVYIMFGGSLPNVIGACIYIILATPIAAIIIPPLGLIYFFVQRFYVASSR 1135
Qy 1139 QLRKLESYSRSPVYSHNETLLGVSVIRAFEBRERFIHOSDKVDENOKAYPSIYANRW 1198
Db 1136 QLRKLESYSRSPVYSHNETLLGVSVIRAFEBRERFIHOSDKVDENOKAYPSIYANRW 1195
Qy 1199 LAVRLCEVGNCLVLAALFAVISRHSLSAGLVGLSVYSISQVTTYINMLVMSSEMETNI 1258

Db 1196 LAVALGECNCIVLFAALFAVISRSHSLAGLVGHSVSYSLQITAYLVMLVMSSEMETNI 1255
Qy 1259 VAVRRLKEYSETEKEAWMOIOETAPSSWPQVGEVERNRNYCLARERDLPFLRINTANTIN 1318
Db 1256 VAVRRLKEYSETEKEAWMOIOETAPSSWPQVGEVERNRNYCLARERDLPFLRINTANTIN 1315
Qy 1319 GGEKVGIVGRTAGKSSLTGLFRINSAGEEIIIDGINIAKIGLHDLRFKTTIIPDDPV 1378
Db 1316 GGEKVGIVGRTAGKSSLTGLFRINSAGEEIIIDGINIAKIGLHDLRFKTTIIPDDPV 1375
Qy 1379 LFGSGLRMTNIDPFSQYSDEEVMATSLAHLKDPVSALPDKLHDCAGGENTSVGOROLV 1438
Db 1376 LFGSGLRMTNIDPFSQYSDEEVMATSLAHLKDPVSALPDKLHDCAGGENTSVGOROLV 1435
Qy 1439 CLARALLRKTILVLDATAVVDLETDLLIOSTIRTOPEDCCTVTIARLNTIMDYRVI 1498
Db 1436 CLARALLRKTILVLDATAVVDLETDLLIOSTIRTOPEDCCTVTIARLNTIMDYRVI 1495
Qy 1499 VLDKGEIOEYGAPSDLLQORGLFYSMAKDGVLV 1531
Db 1496 VLDKGEIOEYGAPSDLLQORGLFYSMAKDGVLV 1528

RESULT 15
US-10-618-281-63
; Sequence 63, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618, 281
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 1303
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-618-281-63

Query Match 61.3%; Score 6068.5; DB 5; Length 1303;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1229; Conservative 17; Mismatches 57; Indels 207; Gaps 10;

Qy 22 MNTSNPFTKCFONTVLVWVPCFYLMACPFYFLYSRHDGYIOMTPLNKTATAGFL 81
Db 1 MNTSNPFTKCFONTVLVWVPCFYLMACPFYFLYSRHDGYIOMTPLNKTATAGFL 60
Qy 82 WIYCMADLFYSFMRSGIGTLAPVFLVSPULLGTTLLATFLIQLERRKVOSSGIMLTF 141
Db 61 WIYCMADLFYSFMRSGIGTLAPVFLVSPULLGTTLLATFLIQLERRKVOSSGIMLTF 120
Qy 142 WLVALCALALILSKIMTAKEDAQVDLFRDITFYVYFSLLIQVLVSCSDSPLEFSET 201
Db 121 WLVALCALALILSKIMTAKEDAQVDLFRDITFYVYFSLLIQVLVSCSDSPLEFSET 177
Qy 202 IHDNPNCPRESSASFLSHITFMWITGLIVRGYRQPLEGSDLSINKEDTSBOVVPVLVKN 261
Db 178 IHDNPNCPRESSASFLSHITFMWITGLIVRGYRQPLEGSDLSINKEDTSBOVVPVLVKN 237
Qy 262 KKECATRKQPVAVVYVSKPAPQKSSKVDANEVVALIVKSPQKMNPSLKVLYKTF 321
Db 238 KKECATRKQPVAVVYVSKPAPQKSSKVDANEVVALIVKSPQKMNPSLKVLYKTF 271
Qy 322 GPYFLMSFFPKAIDMLMFGSPOLLKLLIFVNDTRKAPDWQGYFYVTLFVTACLOTLVL 381

Db 272 QALSL-----LCRLLIKFVNDTRKAPDWQGYFYVTLFVTACLOTLVL 314
Qy 382 HOYFHICFSGMRKITAIVCAVVRKALVITNSARKSSTVGEIYNLSVDQORFMDLATYI 441
Db 315 HOYFHICFSGMRKITAIVCAVVRKALVITNSARKSSTVGEIYNLSVDQORFMDLATYI 374
Qy 442 NMTWSAPLOVITLALYLMNLGSPVLAVGAVVWMLMPVNAVMMKTTQVAVHAKSDNR 501
Db 375 NMTWSAPLOVITLALYLMNLGSPVLAVGAVVWMLMPVNAVMMKTTQVAVHAKSDNR 416
Qy 502 IKLMNBLNGIKYIKLYAMELAFKDKVLAIROBELKYLKKSAYLSAVGTFTWCTPELVA 561
Db 417 IKLMNBLNGIKYIKLYAMELAFKDKVLAIROBELKYLKKSAYLSAVGTFTWCTPELVA 476
Qy 562 LCTPAAVYVITDENNIIDAOCTAFVSLALFNILRPPLNLPVAVISIVQASVSLRLRIFLS 621
Db 477 LCTPAAVYVITDENNIIDAOCTAFVSLALFNILRPPLNLPVAVISIVQASVSLRLRIFLS 528
Qy 622 HEELEPDSIERRPVKKGGGNSITVRNATFTARSDEPTLNGITFSPGALVAVVGQV 681
Db 529 GATSENGPMGSRFRKGG-----TROASFSVAEPVLCRFSITFSPGALVAVVGQV 581
Qy 682 CGKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAMIQNDLRENILFGCOLLEPPYRSVI 741
Db 582 CGKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAMIQNDLRENILFGCOLLEPPYRSVI 641
Qy 742 QACALLPDLIELPSGRTTEIGKGVNLGGQKORVSLARAIVSNADITYLFDPLSAVDH 801
Db 642 QACALLPDLIELPSGRTTEIGKGVNLGGQKORVSLARAIVSNADITYLFDPLSAVDH 701
Qy 802 VGHIFENYVIGPMGMLKNTKRLIVTHSMSTLPQVDVIVWSGKISEMSYQELIARDA 861
Db 702 VGHIFENYVIGPMGMLKNTKRLIVTHSMSTLPQVDVIVWSGKISEMSYQELIARDA 724
Qy 862 FAELRTVASTBOEQAENEVGVGSGPKAKOMENGMVTDASAGOLQRLSSSSYS 921
Db 725 -----S 725
Qy 922 GDISRHNSTABLQAKAKEETWKLMEADKAQTGVKLSVYDYNKAIQLFISFLIFL 981
Db 726 CDL-----QVKLSVYDYNKAIQLFISFLIFL 753
Qy 982 FMCNHHVSAASNWLSLMTDDPIVNGTOERTKRLSVYALGISQIAPFGYSMAVSIIG 1041
Db 754 FMCNHHVSAASNWLSLMTDDPIVNGTOERTKRLSVYALGISQIAPFGYSMAVSIIG 813
Qy 1042 ILASRCIAVDLHSLRSPMSFPERTPSGNLVNRFSKELDYVDSMPEYIKMFMGSLFNV 1101
Db 814 ILASRCIAVDLHSLRSPMSFPERTPSGNLVNRFSKELDYVDSMPEYIKMFMGSLFNV 873
Qy 1102 IGACIVILLATPIAIIIPPLGIYFPVORFYVASSROLKRLSVSRSPYSHNETLIG 1161
Db 874 IGACIVILLATPIAIIIPPLGIYFPVORFYVASSROLKRLSVSRSPYSHNETLIG 933
Qy 1162 VSVYIRAFEOERTIHOSDLKVDENQAKYIPSIYANMWLAVERLECGVNCIVLPAALRAVIS 1221
Db 934 VSVYIRAFEOERTIHOSDLKVDENQAKYIPSIYANMWLAVERLECGVNCIVLPAALRAVIS 993
Qy 1222 RHLISAGLVGLSVSYLOVTTYLNLVMSSEMETNI VVERLKEYSETEKEAWMOIOET 1281
Db 994 RHLISAGLVGLSVSYLOVTTYLNLVMSSEMETNI VVERLKEYSETEKEAWMOIOET 1053
Qy 1282 APPSSWPQVGRVFNRYCLRYREDLDFVLRHINVTINGEKVIGVGTAGKSSLTGLF 1341
Db 1054 APPSSWPQVGRVFNRYCLRYREDLDFVLRHINVTINGEKVIGVGTAGKSSLTGLF 1113
Qy 1342 RINESAGEEIIIDGINIAKIGLHDLRFKTTIIPDDPVLSGSLRMTNIDPFSQYSDEEVM 1401
Db 1114 RINESAGEEIIIDGINIAKIGLHDLRFKTTIIPDDPVLSGSLRMTNIDPFSQYSDEEVM 1173
Qy 1402 SLELAHKDFVSALPDKLHDCAGGENTSVGOROLVCLARALLRKTILVLDATAVVD 1461

Db	1174	SLELAHLKDFVSALPDKLDHECAEGGENLSVGORQVCLARALLRKTILVDEATAVD	1233
Qy	1462	LETDLLIOSTIRIQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQORGLF	1521
Db	1234	LETDLLIOSTIRIQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQORGLF	1293
Qy	1522	YSMAKDAGLV	1531
Db	1294	YSMAKDAGLV	1303

Search completed: December 15, 2005, 15:47:08
Job time : 171.832 secs

QY 855 LIAADGAPAEFLRTYASTEQDABENGVTGSGPKAKOMENGMTVTDSAGK-QIQR 912
 Db 360 LLOKRGKLVCLL-----DQARQPDGRG-EGETEGTSTKDPRG-----TSAGRPBELRR 407
 QY 913 QLSSSSYSGDISHHNSTAELOKABAKEET---WKLMEADKACQGVCLSYWMDYMK 968
 Db 408 ERSRKS-----VPEKDRITSEAO-TEVPLDDPDRAGWAGK-DSIQYGRKATVAILAYLR 460
 QY 969 AIGLFISFLFEMCNHVSALASNYWLSLWTDDEPIVNGTOEHTKVLRSVYGALGISQGI 1028
 Db 461 AVGPPLCYALFELCQOVASFRCGYWLSLWADDPAGVGGQQTQALRGIGLGCQAI 520
 QY 1029 AVFEGYNAVSIGILASCLHVDLHSLRSPMSFEETPESGMLVNRSEKLDYDSMIP 1088
 Db 521 GLFASMAVLLGGARASRLFORLMDVRSPISEFFERTPIGHLNRFESKEDTVVDVIP 580
 QY 1089 EVIMFGMSLFNVGACIVILLATPPIAIIIPGLIYFVQRFVAVASSRDLKLESVR 1148
 Db 581 DKLRSLMYAFGLFESLVAVATPLATVAILPLFLYAGFOSLYVSSCQURLESASY 640
 QY 1149 SPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLAVLCEVGN 1208
 Db 641 SSVCSHMAETFGGTVVRAFRTOAPFVAGNNARVDESQRISFPRLVADRWLAAVETLGN 700
 QY 1209 CIYLPALFANISKHSLSAGVGSYSLOVTTYTLNVLNMSSEMETNYIYVERLKEYS 1268
 Db 701 GLVPAATCAVLSKXHSAGLVGSVSALQVTQLQVVRNMDIENSIVSVERMODYA 760
 QY 1269 ETEKAPMOIOETAPSPSPPOVGRVPEFRNYCGLAREDDFVLRIHNTVITNGEKYIGR 1328
 Db 761 WTPKEAPRRLPTCAQPPWPGGQIEFRDFGLRIHPELPLAVQGVSEKIHGEKGIYGR 820
 QY 1329 TGAKSSLTLLGFINSEAGEIITDGINAKIGHLDLRFKITIIPQDPVLFSSGLRNL 1388
 Db 821 TGAKSSLTASGLNLQGEABEGIMWDGPIAHVGLHSLRSISIRPPIIFPSGLRNL 880
 QY 1389 DPEQYSDDEWTSLELAKDQFVSALPDKLHCEAGEGENLVSQOLVCLARALRKT 1448
 Db 881 DLLOHSEDEAIWALETVQLKALVASLPGOLQYKCADGEBLSVQKOLLCLARALRKT 940
 QY 1449 KIIVLDEKTAIVDELTDDLGISTIRTOFEDCTVLTARLMTIMTYTVIYLDGELOEY 1508
 Db 941 QILLDEKTAIVDPELTLOMOMLGSWFAQCTLLIARLSVMDCAVWLMDKQVAVES 1000
 QY 1509 GAPSDDLQORGLFYFSMAKDAGLV 1531
 Db 1001 GSPAQLLAQKGLFYRLAQESGLV 1023

RESULT 2
 US-11-090-439-24
 ; Sequence 24, Application US/11090439
 ; Publication No. US20050266442A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Squillace, Rachel
 ; APPLICANT: Weiner, Michael P.
 ; TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
 ; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
 ; FILE REFERENCE: 24318-502
 ; CURRENT APPLICATION NUMBER: US/11/090,439
 ; CURRENT FILING DATE: 2005-03-25
 ; PRIOR APPLICATION NUMBER: 60/556,344
 ; PRIOR FILING DATE: 2004-03-25
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 24
 ; LENGTH: 1581
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-090-439-24

Query Match 21.4%; Score 2115; DB 7; Length 1581;
 Best Local Similarity 31.9%; Pred. No. 3,8e-147;

Matches 527; Conservative 305; Mismatches 564; Indels 256; Gaps 37;
 QY 32 CFQNTVLVWPCFYLMACFPFYP-LYLRHDSRGYIOMTPLNKTALGPF---LIMVICW 86
 Db 26 CFVDALNVVPHVFLFLFTPLFLIGWSQSSKVIHHS-----TWLHPQHLRWILTF 79
 QY 87 ADLFYSFWEBSRGITFLAV-----FLVSPF---LIGITTLATFLIQLERRGVSSG 136
 Db 80 MLFLVLCESIAEGILSDGVESHHLIYMPAGAFMAVIVSV--YYHNIETSNFPLLI 137
 QY 137 IMLFVVALVICALALIRSKIMTALKEDAOVDLFRDITF-----VYPSLLIQLVLS 189
 Db 138 ALVYMTLAFI-----TKIKFVKFLDHAIGESQURFLCTGLVLYIGKMLIVE--VN 188
 QY 190 CFSRSPFLFSETIHPNCPES-----SASFISRTFWMITGLYVRGROPLEG 238
 Db 189 VIRRRRIYIFRTPEVRP-PEDLODLGVRLFQPVNLKSTYMNNAFITAHKPI-- 245
 QY 239 SDSLANKEDTSEOVVVLK--NMKKECAKTRKQVVKVYSSKDPAPKSSVDANE 295
 Db 246 -DLRAIGK-----LEIARALNTYORLC-----E 268
 QY 296 EVELIYKSPQ-KEMNPSLFFVLKYKTFGYFLMSFFPKAHDLMFSGPOLIKLIRKV- 353
 Db 269 AFDQVRKDIOGTGAPALWALSHPGRVLVSTSTRILADLGFAGPLCTFGIVPHIG 328
 QY 354 -NDTKAPDQ-----GYFYVLLEVTAQIOTVLAHOYFHICFVSGMRK 396
 Db 329 KENVPPQPKQFLGVFVSSQEPFLANNVYLAIVLLFLLRITQFLQASYYVAIEGTINLR 388
 QY 397 TAVIGAYYRKA--LVITNSARKSSTVEIYNLMSVDQREMDLATTYINMTASAPLOYILA 454
 Db 389 GAIOTKLYNKMHLSTNLSGEBWTAQICNLVAIDTNQMLFEPFLPNLMAVMOQIIVG 448
 QY 455 LYILMLNGBSVLAGVAVMVLMPVNAVMAKKTQVQVAHKSQNDIKLNEITLIGKV 514
 Db 449 VILLYYLLGVSALLGAIVTILALPVQYFVATKLSQAPRSTLEYSNERKQJNEMLRGKL 508
 QY 515 LKLYAMELAFKDKVLAIROBELKYLKKSAYLSAVGFTWVCTPLVALCTEPAVVYITDEN 574
 Db 509 LKLYAMENIIFPTRETRRREKMSLRAFAITYSISIMNTAIPFAVILITVGHVSFPKE 568
 QY 575 NILDAQFVSLALFNLRFPLNLIPNVISITVOASVSLKRLIFLSHBELEPDSI--- 630
 Db 569 ADFSPVAFASLIFLILVPLFLSSVSVSTVALVSVQLSSEFLSABEIREBOCAPHE 628
 QY 631 -----ERRPVK-----DGGGN-SITVANAFTFW 653
 Db 629 PTPQSPASKYQAVPLRVNRRPRAREDCRGITGLQSLVPSADGDADNCCVQIMGYFTW 688
 QY 654 ARSDPTLNGITFSPREGALVAVVGVGCGKSSLSALLAEMDRYEG----- 700
 Db 689 TPDDIPTLSNTITIPRGQLTMIVGVGCGSSLLLAALGMOQVSGAVFWSLDPSEIG 748
 QY 701 -----HVAIKGSVAVYPQAMQIONDSLRENILFGQLEBPYRSVIOACALL 747
 Db 749 EDPSERETATDLDIRRGPVAVYASQFWPLNATVEENIIFESPFRKORYGVATEASLO 808
 QY 748 PDLEILBSGRTETEGEGVNLSSGQOKQVSLARAVNSADLYLDDPLSANDAVGKHTF 807
 Db 809 PDIDILPHGDQTOIGEGINISGGQORISVARALYOHANVFLDDPSSALDIHLSHLM 868
 QY 808 ENVIGPKGMLK-----NKTRLVTHNSYLPQVNVITVMSGKLSKMSGYOELLARDAF 862
 Db 869 Q-----AGILELRLDDKRTVVLVTHKLOYLPHADWITLAKDGTIQREBITLKDPORSEQL 923
 QY 863 AEPLRTYASTEQDABENGVTGSGPKAKOMENGMTVDSAGKQLOROLSSSSSYSG 922
 Db 924 FEHWKTLMN-RQDDELEKETV-----ERKATEEPQG-----LSRAMSSRDGL- 965
 QY 923 DISRHNSTAELOKABAKEETWKLMEADK-----QTGVKLSVYMDYKAIGLFLISFL 977
 Db 966 -----LQDEEEEBEBAESEDNLSMLHORARLIPWPAKAKYLSAGIILLST 1014

Qy	1084	USMIPEVIMKMGSLFENVIGCIYVLLATPTLAI111PLGLIYFVQSEFYVASSQOLKRL	1143
Db	1132	DQIHPISTIECISRSTLLCVSALAVSYTPFVLALPLAIVCYTQKRFVASDLOOL	1191
Qy	1144	ESVRSPLYSHFNETLLGVSVIRAFEBQRFTHQSLKVDENOKAYPSIYANRKLAVRL	1203
Db	1192	DDTQPLPILSHPAETVEGILTIRAFRYEARFOQKLEVTDSNNIASLPLTANRRLLEVBM	1251
Qy	1204	-----ECVGNCIYLFPAALPRAVIS--RHSLSAGLVGSYSLSLOVTTYLANVAMSEM	1255
Db	1252	ATPLPEQYIACAVVLLIAVTSINSLSRRELSAGLVGGLTYALAVSYLMMVMVNLADM	1311
Qy	1255	ETNIYAVERLKEYSETEKE-----APMOIETAPSSWPQYGRVGFNRYCRLRYEDDDF	1308
Db	1312	ELQGVAVRHHGLKLTAEASVEGLAPSLI-----PKXWPOGKQIQIQLSTRYDSSLSK	1366
Qy	1309	VLRHINTVINGEKGVIVGRTGAGKSSLTLLGLFRINESABEHI1IDGINIAKIGHADLRF	1366
Db	1367	VLKHVNALISPGQKIGICGRTSGGKSSLSLAFPRVWDPEGH11IDGIDIALPLHTRS	1428
Qy	1369	KITTIPOBPVLFSGSLKMNLDPOFSQYSEBWTSLSLHLKOPFASLDPKLDHECABEGE	1428
Db	1427	RLSIILOPVPFLPSGTFIRPNLDBERKCSSTLMEALH1QOLKVVRALPQGDALITTEGSE	1486
Qy	1429	NLSVGQROLVCLARALLKRLKYLVLDEATAVADLETDDLQSTIRTOPEDCITVLTIAHRL	1486
Db	1487	NPSGQGFOLFCIARAFVAKTSIFIDEVNTASIDMATEN1LQCVMTAFADRTVITIAHRY	1546
Qy	1489	NTIMDYTRYVULDKGEIOEYAPSPULLOOR--GLFYSMAK	1526
Db	1547	HTIISADLVIVLKRGAILEPDKPEKLLSRKQSVESAFVR	1585

```

RESULT 4
US-11-090-439-26
; Sequence 26, Application US/11090439
; Publication No. US20050266442A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 1581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-26

```

Query Match	21.2%;	Score 2100;	DB 7;	Length 1581;
Best Local Similarity	31.7%;	Pred. No. 4,8e-146;		
Matches 524;	Conservative 306;	Mismatches 567;	Indels 254;	Gaps 37;

QY 32 CFQNTVLVWVPCFYLMACPPFFL--YLSRHDRGYQIMTPRLNKTALGF---LIMVWC 86
 Db 26 CFVDALINAVPHVFFLLIFTEPILFIQWGSQSSKVHIHS-----TWLHPHGHNLMILTE 79
 QY 87 ADLFISFMRKSGITLAPV-----FLVSPT---LLGITTLAFLIQLERKKVQSSG 136
 Db 80 MLFLVACEIAGSILSDGVTSHHLLHLYMPAGMAFMAATVSV--YHNHIFSNPKILLI 137
 QY 137 IMLTFWLVALVCAALILSKINTALKEDAQVDLPFDITF-----YTFSLILLIQVLVS 189
 Db 138 ALLVVTYLAFL-----TKTKIFVKLLDHAIGFQLRCLGLGLVILVIGMLLVE--VN 188
 QY 190 CFSDBSPLFSEF---IHDPNCPBBSASTL-----SRITWMTTGLIVRGYRQPLBGS 239

D	b	189	VIRRRRYEFFETPREVXPREDLODLGRFLDPFNLSKGYWMMNAFKTAHKKPI---	245
Q	y	240	DLMSLNKEDTSEQVPPVLVK---NMKCECAKTRKQPVKVYSSKDPAQKESKYDANE	296
D	b	246	DLRAIGK-----LPIYMRALTYVORLC-----EA	269
Q	y	297	VEALIVSPQ-KEMNPSLFKLYTTCGYPLMSFPFKAIHDMMSGPOILKILIKFY--	353
D	b	270	FDQVORXDIQGTQGARAIWQLSHAFGRRLVLSFTFRLADLLGAPAGLCTFGIVDHGK	329
Q	y	354	-NDRKAPDMQ-----GYFYVTLTPVTAQLOTLVHQHFIFGSGMRKT	397
D	b	330	ENDVPOPEPTQGLGYVPSSQGFELNAYVNLTLFLALLQGTFLQASYYVAIETGINRG	389
Q	y	398	AVIGAVYRKA--LVIITSARKSFTVGEIWNIMSDVQRFMDLATYINNIMSAPLOVIAL	455
D	b	390	AIQTKTYKIMHLSNLSMGEWTAGQCNLVAIDTNQIMWFPEFCPLTAMAPQIIVGV	449
Q	y	456	YLLMLNLGPSVLAGVAVWVLPVNNAVAMTKTYQYAHMSKNRITLMLNEILNGIKVL	515
D	b	450	ILYYIIGVSLTGAAYIILAPQYFATYLSQAQSTTEYSNERLKQTNEMTRGIKLL	509
Q	y	516	KLYAMELAFKQYALROEBLKVKLSAYSAVGTFFWCPPLVALCTFAVYTYIDENN	575
D	b	510	KLYAMENIFRRVETRKREKMTSLRAIAYISISIFMNTAIPLANVLTTFGHVYSFPEA	569
Q	y	576	ILDAQAFVSLALFNILRFPNLIMPVYISSIVQASVSLKRLRIFLSHEELEDST----	630
D	b	570	DEPSVAPASLSLPHILVTPLFLSSVYRSTVKALVSQKLEFSLSAEIREEQAPHEP	629
Q	y	631	-----ERRPVK-----DGGTN-SIVYRNATFPWA	654
D	b	630	TPQGBASKYQAVPLRVNRRKPAEDCGLTGPLQSLVPSADGNCCVQIMGIFYTWT	689
Q	y	655	RSDPELTNGIFESIPEGALVAVWQOVGCGKSSLSALLAENDKYEG-----	700
D	b	650	PDGIPTLSSNTITRIPRGOLTMIVQOVGCGKSSLLALGEOXKYSGAVFMSLSDEIGE	749
Q	y	701	-----HVAIKGSVAVYVPQAWIIONDSIRENITLPGCOBEPYRSVIOACALLP	748
D	b	750	DSPERERATDLDLRKRGPVAVYASQKPLNATAYENIIFESPKNQKQKMYIEACSLQP	809
Q	y	749	DLEIIPSGDRFIEIGEKVNSGGQKORVSLARAYSNADYTLFDDPLSAVDYAHGKHIFE	808
D	b	810	DIDILPHEDQOIGERGINSGGGRORISVARALYOHANVFLDDPFALDIHLSDHMQ	869
Q	y	809	NVIGKGMK-----NKTRIIVTSMSTLPQVDYIYVMSGGKISGMSGYOELBLARDAFA	863
D	b	870	-----AGLEILLRDDRKTIVLVTHKQVLPFADMIIAKQDSTIOREGLTQPOFQSECOLF	924
Q	y	864	EFLRTYASTEOEODAEENGVTGVSGBPGKEAKOMENGLVYTSAGKQLOROLSSSSYSGD	923
D	b	925	EHWKTIWM-RQDQLEKETV-----EKKAIEPQO-----LSRAMSBDGL---	965
Q	y	924	ISRRHNSTAELQKAEKKEETWKLMEADK-----QTGQVLSYYWDMKAIGLFISLS	978
D	b	966	-----IQDEEBEEFAESEDNDNLSMLHQRAEIPWRACAKYLSAGIILSLTL	1015
Q	y	979	IFLEPMCHVSLASNYMLSLMTDDPIV-----NGQOHT-----KXLSYYGALGI	1024
D	b	1016	VFSQLKTMVVAIDYMLAKWTDLSLTTPAARNCSLSQECTLDQTYVAMFYVLCSTGI	1075
Q	y	1025	SOGIAVFGYMAVNSIGGILASRCJLHVDLSILRSPMSFERTSPSGNLVNRFSKELDTVD	1088
D	b	1076	---VLCIVTSYTVEMTGKXAKRLHRSLLNNIILAPMRFFETPLGSLINFFSDCNTID	1133
Q	y	1085	SMIEPVIMFGSLFNIVGACIVILLATPIAAIIIPPLGIYFFQRYFVASRQKLE	1144
D	b	1133	QHIBTEECTSRSTLLCVSALAVISYVPEVLVALLPLAVCVFYQKERYVASRDLQOLD	1192
Q	y	1145	SVSRSPVYSHNETLGVSVIRABEQRFLHOSDKVDENQAKAYBSIVANRMLAVLE	1204
D	b	1193	DTQOPLLSHAEVTEGJLTIRAREYERFOOKLEYEDSNINISLFLTANRMLTEVME	1253

Qy	1205	CVGNQIVTFFAALFAVIS--RHSLSAGLVGLSVSYSLQVTTYLNTMLVRMSSEKTEINI VAVE	1265
Db	1253	YTGACVVLIAAVTSTISNSLHRLSLGVLGLTYLALMNSNTLMMVNRVLADMELOLGAVK	1312
Qy	1263	RLKEYSETEKE-----APWOIQETAPSPSWQVGRVFRNYCLRYREDLDFVLRHINVT	1316
Db	1313	RHGLLTKTEAESYEGELALPSPIL-----PKMWDQCKIQLQNLVSRYDSSLRKVLGHVVAL	1367
Qy	1317	INGGRKQIVTERTGKSGSLTGLFRINBSAGELIIDGINAKIGLHDLRPKTITIFOD	1378
Db	1368	ISPQKIGICGATGSGKSSFSFLAFPRMVDTEFGHIIIDIGIDIRKILPLATLPSRSLIIOD	1422
Qy	1377	PVLFSGSLRMNLDPPFSOXSDDEEVMVTSLELAHLKDIFYSALPDKLDHECAGGGENLVSQRO	1436
Db	1428	PVLBSGTRFRLNLPDRKCSDSLTMALALIAQLKLVYKALPGGLDAITTEGGENFSQGRQ	1488
Qy	1437	LVCLARALLRKTKILVLDEATAVADLETDDLIQSTIRFOPEDCVLTIAHRLNTIMDYTR	1496
Db	1488	LFCLARAFVRKRTSIFIMDEATASIMATENILQKVMTAFADRTVVTIAHVRHTLSADL	1547
Qy	1497	VIVLDKGEIOEYGAPSDILQQR-GLFYFSMAK 1526	
Db	1548	VIVLRGALILEFDKPEKLLSRDVSFAFSVR 1578	
RESULT 5			
US-10-995-561-526			
; Sequence 526, Application US/10995561			
; Publication No. US20050272054A1			
; GENERAL INFORMATION:			
; APPLICANT: CARGILL, Michele et al.			
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH			
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF			
; TITLE OF INVENTION: DETECTION AND USES THEREOF			
; FILE REFERENCE: CLO001559			
; CURRENT APPLICATION NUMBER: US/10/995,561			
; CURRENT FILING DATE: 2004-11-24			
; NUMBER OF SEQ ID NOS: 85702			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 526			
; LENGTH: 1404			
; TYPE: PR1			
; ORGANISM1: Homo sapiens			
US-10-995-561-526			
Query Match 13.5%; Score 1341.5; DB 6; Length 1404;			
Beet Local Similarity 27.8%; Pred. No. 2.2e-90;			
Matches 369; Conservative 242; Mismatches 473; Indels 243; Gaps 34;			
Qy	32	CFONTVLVWVOCFYLMACFPYFL-YLRRHNRGVIQMTPLNKTKTALGF-----LIMTICW 86	
Db	26	CFVDALNVVPRVFLITFTFFLITFTIGWSQSSKVHIHS-----TWLHPGCHNLRLWLTPE 79	
Qy	87	ADLFYSFEMERSRGIFLAPV-----FLVSPPT-----LIGITTLATFLIQLERRKGVQSSG 136	
Db	80	MLLFVLVCEIAEGILSDGVESHHLHLVMPAGMAFMAVTSVV--YHNITSNPPKLLI 137	
Qy	137	IMLTFVLVALCALALRSKMTALXEDAOVDLPDITF-----YVFSILLIQLVLS 189	
Db	138	ALLVYMTLALF-----TKTKIKVFKVLDHAIIGSFQLFCUTGLVLIVYGMILLIVE--VN 188	
Qy	190	CFDSRPLFSFETIHDPNCPES-----SASFISRTFFMWITGLIVRGROPLEG 238	
Db	189	VIRVRRIYFFFTPREVKE--PEDDLDLGVRFLQDPYNLLSKGTYYMMNAFITAKKPT-- 245	
Qy	239	SDLSLAKEDTSBOVVPVLVK--NMKKECAKTRQPVKVVYSSKDPAPQYESSKVDANE 295	
Db	246	-DELAIGK-----LPIAMRALNTYQRLC-----E 268	
Qy	296	EEVALIVSPQ--KEENPSLEFVLVKTGPRPYFLMSFFFAIHDLMFSGPOLIKLLIKRV- 353	
Db	269	AFDQVQRKDIOGTQGARAIWQLSHAFERRVLVLSFTFRIADLDLGFAGPRLIFGIVDHLG 328	

```

0Y 354 -NDPKADMQ-----GFEYTLTVTCTQTLVHQHYHCFVSGMRJK 396
Db 329 KENDFOEKTQFLGYFWVSSOEFLANAYVAVLTLALLQRTFFLOASTYVVALETGINR 388
0Y 397 TAVIGAVVRK--LVITNSARKSSTVGEIWNLSYDAORFMDLATYINNMISAPLOVILA 454
Db 389 GAIQKIKYKIMHISTNLSMGEMTAGQICNVLADTQNMFFFLCPRLMAMPVOIYIG 448
0Y 445 LYLMLNIGPSVLAVAVMLVMPVNAVMAKTKTYQVAHMSKONRITKANELLNGIRY 514
Db 449 VILYYILIGVSALIGAAVILILAPQYEVATRLSOQORSTLEYSNERLQOTEMELRGIKL 508
0Y 515 LKJWAMELAFDQKVALROEBELKVLKKSAYLVAOTFFWCVCPVLVALCTFAVYVYIDEN 574
Db 509 LKJWAMEI FETIRVETTRKEMTSLRAFALYTSISI FMTALPAAVLTFVGHVSFEKE 568
0Y 575 NILDQAOTFVSLAENILRFPNLILPMYISSIVQASVSLKRLRIFLSHEBELPPDSI---- 630
Db 569 ADFSPSVAFASLSPLHILVTLPLFLSSVVRSTVKALVSGKLSEFLSSAEINEBOCAPIH 628
0Y 631 -----ERRPVK-----DEGCTN-STVKNATPFW 653
Db 629 PTPQGPASKYOAVPLRVLVNRKRPAEEDCRGLTGPLQSLVPSADGADNCCVQIMGGYFTW 688
0Y 654 ARSDPPTNGITFESIPEGALVAVYQVQCGKSSILLSALAEADKYEG----- 700
Db 689 TPDGIFLISNITIRIPRGQLTWYIGQVQCGSSILLALGEMOKSGAVFWSSLPDSEIG 748
0Y 701 -----HVAIKGVAAYVPOQAMIQNDSIRENILRFGCOLEEPYRSVIOACALL 747
Db 749 EDPSPERETADDIRKGPVAVASQKRWMLNATVEENIIFESPENKQRYKNVIEACSLQ 808
0Y 748 PDLEILPBGDTEIGEKEVNISSGGCKQVSLARAVYSNADYTLFDDPLSAVDAYHKAIF 807
Db 809 PDIDILPHGDQTOIGERGINSSGGQRORISVARALYOHANVFLDDPFSAIDIHSDHLM 868
0Y 808 ENVIKPKMKL-----NKTRILVTHSMVYTLPOVDYIVYMSGKISSEMGSYOBELARDGAF 862
Db 869 Q-----AGILELRDDKRTVVLVTHKLOYLPHADNIIAMKOSTIQREGLTKOPORSQOL 923
0Y 863 AEFRLTYASTBOBDAENGVTVGSGPKEAKOMENGMVLVTDSSAKQOLQROLSSSSYSBG 922
Db 924 FEHWKTLMN-RQODELEKETV-----ERKATEPQG-----LSRAMSRDGL-- 965
0Y 923 DISRHNSTAELOKRAKEETWKLMEADKA-----QTCQVYKLSYVWDYMKALGLEIFL 977
Db 966 -----LOEEEBEEBAASEEDNDNLSMLHQABEIPWACAKYLSAGAILLSTL 101
0Y 978 SIFLFMCNHSVALSNMYLSMTDDPIV-----NGNOEHT-----KVLASYGALG 102
Db 1015 LVFEOQLKHMVLVADIVLAKTTBSALTLTPAARCSISQEBCTLDQTYAMAVFLVCSIG 107
0Y 1024 ISOGIAVFGVSMAYSIGGILASRCLVDLHLSILRSPMSFFERTPSGNLVNRFSELDTV 108
Db 1075 I---VLCJWVS VTYEMWGLKVAKRLHRSILNRITLAPRPFETTPGLSILNRFSSDCMTI 113
0Y 1084 DSMPEVYKMGSGSLFNVIQACIYIILATPLAAIIPPLGLIYFVQRFYVASSQOLKRL 114
Db 1132 DQHPSTLECSRSTLVCVSLAAIISYVTPVELVALLEPLAIVCEYIQCFYFVASDLOOL 119
0Y 1144 ESVSRPSVYSHFNELLGVSYIRAFEEBOERFLHOSDLKVDENOKKYPSIYVNMRLAVRL 120
Db 1192 DDITQOLPLSHFAETVEGLTIRAFRYEARQOKLEBTDSNNIYASLFLTAANRWLEVRM 125
0Y 1204 ECVGNCTI 1210
Db 1252 -AIRCM 1257

```

```
Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 969
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-969

Query Match          9.8%; Score 969; DB 6; Length 410;
Best Local Similarity 50.8%; Pred. No. 8e-64;
Matches 197; Conservative 66; Mismatches 121; Indels 4; Gaps 2;

QY 495 MRSKDNRIKLNELINGIKVLKLVAMELAFKDKVLAIRQSEELKVLKKSAYLSAVGTFTWV 554
DB 1 MRQDSRAKLTSSILRNKSKITKFKHMEGAFIDRVLGITGDELGLRTSGLLPFSVLSVFQ 60
QY 555 CTPLVVALCTFAVYVYITDENNILDAQTAFVSLAFENILRFPNLIPWVSIIVQASVSLK 614
DB 61 VSTFLVAVFAVHTLVAE-NANNAEKAFVTLTVNLINKAQAFLEFHSILVQARVGF 119
QY 615 RLRTFLSHLEELPDSIERPKYDGGGNTSITVRANFTFWASDPTPLNGITFSPLEGALV 674
DB 120 RLTVFLCLEEVDPGVAVSSSGSAAGKDCITIQSTFAMSOEPECLHRIMLVPOGCLL 179
QY 675 AVGVGVGGKSSLSALLAEKDVKEGHAIVKGVAVYVPOQAMIONDSRENILGSCOLEE 734
DB 180 AVGVGVGGKSSLSALLAEKDVKEGHAIVKGVAVYVPOQAMIONDSRENILGSCOLEE 734
QY 735 PYRVSVIQACALLDLPLBILPSGDRTEIEGKVNLSGGOKORVSLARAVYSNADYILPDDP 794
DB 240 PWLERVLEACALQPDVDSFPBGHITSIGEGMNLSSGKQSLSLARAVYRKAAYLLDDP 299
QY 795 LSAVDAAVGHKIFENVIGPKMKLNKKTITLVTHSKSYLPQVDVITVWGGKISSEKSYOE 854
DB 300 LAALDAHVGHQHFVQVIGPGLLOCTRTIILVTHALHILPOADWITVLNGAIAEWGSYOE 359
QY 855 LLAADGAFAEFLRTVASTEOBDAENG 882
DB 360 LLQKRGALVCL--DQARQPGDRGEGG 384

RESULT 7
US-10-613-744-14
; Sequence 14, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
```

```
SEQ ID NO 14
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-613-744-14

Query Match          8.9%; Score 883.5; DB 6; Length 407;
Best Local Similarity 48.7%; Pred. No. 1.5e-57;
Matches 167; Conservative 74; Mismatches 95; Indels 7; Gaps 4;

QY 1551 LTRAEDEPAEPRY----RTERRARFYSKGCNCAVAKNIREQGRFIDVPTTIVDLKWP 1606
DB 16 LPRQARDL--PRHISRDTRKKIQRYVRKQKCNHGVNREYRYLIDITFTLVLDLKW 74
QY 1607 HTLLIFMTSPFLCSWLLPAMVWMLIAFAHGLDPAEGEGRN-PCVTSIHSFSSAFPLFSIEVQ 1665
DB 75 FNLLIFVWYTVVWLFEGMILWLAIRGMDHDESDPWPVPCVNNLNGVFSAFLFSTETE 134
QY 1666 VTTFGGRMWTEBCEPLAILILIVONIVGLMINALMGLCIPFKTAQAHRAETLIFSQHAV 1725
DB 135 TTIGGYRVVITDKCPREGIILLILQSVLGSIVNAFVGCMEFVKISQPKKRAETLVFSTHAV 194
QY 1726 ITLPHGRLCEPMLRVGDLRKSMIISATIMQVVRKTTSPBGRVVPILHVDIPMENGVGNG 1785
DB 195 ISMDGKLCIMFRVGDLRNSHIVASIRAKLISKQTSBGEFLPLNQTIDINVGYYTGDDR 254
QY 1786 IFVLAPLIYHVIDSNSPLYDLAPSDLHHODLEIIVLBSGVETGTLTTOARTSYLADE 1845
DB 255 LFLVSPILSHIEINQSPFWEISKAOL-PKEELEIVILBSGVATGMTQCARSSYITSE 313
QY 1846 ILWGQRFVPIVAEEDGRYSVDYSKFGNTIKVPIPLCTARQIDE 1888
DB 314 ILMGYRFPLVLTLEDGYEVDYNSFHEVETSTSLAKELAE 356

RESULT 8
US-10-995-561-967
; Sequence 967, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 967
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-967

Query Match          8.1%; Score 799; DB 6; Length 268;
Best Local Similarity 56.3%; Pred. No. 1.3e-51;
Matches 151; Conservative 56; Mismatches 61; Indels 0; Gaps 0;

QY 1264 LKXESETEKAPWQIQETAPSSWPQVGRVVEFRNYCLRYREDLDFVLRHINVTINGEKV 1323
DB 1 MQDYAWTPKEAPWRLPFCQAQPPWPQGGQIEFRDFGLRYRPELPLAVQVGSFKIHAGK 60
QY 1324 GIVRTAGKSSLLGLFRINESAGEGLIIDGINIAKIGLHDLAFKTTIIPODVLPSSGS 1363
DB 61 GIVRTAGKSSLASGLRLQEALEGGIWDGVLAHGLTLSRSLIIPQDPLPFGS 120
QY 1384 LRNMIDPFSSQYSDDEVVTSLELAHKDFVSALPKDLHCECAEGENISVGGQRLVCLARA 1443
DB 121 LRNMIDLQHSDDAIAALETVQLKALVSLPQQLQYKCDKREDDLSVGKQLLCLARA 180
QY 1444 LARKTKILVIDEATAVADLETFDILLQSTIRTOFEDCVLTIAHRLNTIMDYTRVIVLDKG 1503
```


Dh 181 LARKTOILLDEATAAVDPGTELOMQLSGMFAOCTVLLIYHRLRSVMDCARVLVMDKG 240
Qy 1504 EIOEYAPSDLLQORGLFYSMAXDAGIV 1531
Dh 241 QVAGSGSPAQLAQORGLFYRLAQSGLV 268

RESULT 9
US-10-613-744-13

; Sequence 13, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; TITLE OF INVENTION: The Rockefeller University
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: NO PCT/US99/06307
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 391
; TYPE: PR
; ORGANISM: Rattus norvegicus
US-10-613-744-13

Query Match 7.9%; Score 786.5; DB 6; Length 391;
Best Local Similarity 46.2%; Pred. No. 1.9e-50;

Matches 150; Conservative 75; Mismatches 89; Indels 11; Gaps 6;

Qy 1564 RTERRARFVSKKCNVANHKNIREQRF--QDVFTLLVDLKNPHLLIFTMFLGSL 1621
Dh 35 RSRQ-RARLVSKGRCNIEFGNDAQSRFIFVDIMTTLVDLKKRYKMTVITAFILSMF 93
Qy 1622 LFMVWMLIAFAHDL-----APGEGTAVPCYTSIHSSSAFLBEIEQVYTGFGGRWTE 1677
Dh 94 LFGLLWVVAVYVHDLPEFYPPDNR--PCVENINGMTSAFLFLETOVITGYGFRFVTE 151
Qy 1678 ECPALAILLYONIVGMINAIMGCTFMKTAOHRRAETLIPSKHAVITLHGRLEGL 1737
Dh 152 QCATATLFLFOSILGVIINSFMCALLAKISRKPKAKTITFSKNAVISRGGKCLLI 211
Qy 1738 RVGDLRKSMTISATIHQVVRKTSPEGEVVPFLHQVDIPMENGVGNGIFLVAFLIYHV 1797
Dh 212 RVANLRKSLIGSHITVGLTKTITTPGETIITLDQTNINPVVADGNENLFPISPLIYHI 271
Qy 1798 IDSNSPLVDLAPSLDHHODEIIVIEGVVETTGITTOARTSYLADILLMGORFPIVA 1857
Dh 272 IDHNSPFHMAAETL--SQODFELVFLDGVTESTSATCQVRTSYVPEEVLMGYRFPIVS 330
Qy 1858 E-EDGRYSVDYSKRGNTIKVPTPLC 1881
Dh 331 KTKGKTRVDFPHNFGKTVETPHC 355

RESULT 10
US-10-793-626-326
; Sequence 326, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US

; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 326
; LENGTH: 548
; TYPE: PR
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-326

Query Match 5.2%; Score 514.5; DB 6; Length 548;
Best Local Similarity 27.1%; Pred. No. 3.2e-30;
Matches 157; Conservative 110; Mismatches 211; Indels 101; Gaps 16;

Qy 1002 DPTVNGTQHTKRLSYGALGISQGIYAVGYMAVSIIGILASRCLHVDLHSLTSPM 1061
Dh 12 DGVIN--NHSLTNQEKFSHIGVAILGF-----IFLIVRPI 47
Qy 1062 SFPE---RTPSGNLVRFSEKL-----DTVDSMPEYIK-----MFMG 1096
Dh 48 EPIROYLAQMTSKNLIYDQKQNLNHLQALSVREFYANNOYGVISRVINDVEQTKDITLT 107
Qy 1097 SLENVIGACIVILLAPI-----AAIIPPLGL--IYFVQGFYVASSRQLRLE 1144
Dh 108 GLNMIMDCITIIIALISIMFPLDVKLPFAIIFPPFYLIVYFFPGL-----RKLTRVR 162
Qy 1145 SVSRSPYSHNFTLGVSVITRAFEDEGERTHOSDLKVDENQKXYYISYANRLAARLE 1204
Dh 163 SQALAEVQGFLLHERVQMSVYKSFALIEDNAKFD--NHNKQFLQRAFOHTRWNAYSFA 219
Qy 1205 CV-----GNCIVLFAALFVVISRHSLSAGLVGSYSLOVTTYLMVLMSEMETNI 1258
Dh 220 AINTVTDLGLPIIYIGVSYLAIT--GSTITVTLAFCVYLEQLRPLRLVSSFTTLTQSF 278
Qy 1259 VAVERLKEYSBTEKE-----APWQIQETAPSSWPQVQVREVRNYCLARYEDLDFVL 1311
Dh 279 ASMDRVQLMDEDDIDKNGIQAQPIKSK-----QIDLKHSFYNNKEKEVHL 328
Qy 1312 HINVTNGEKRGVIGVRTGKSSLTGLFRINESABGEIITIGINIAKGLHDKKIT 1371
Dh 329 DINTLTNGKGTVAFCVSGGSKSTLINLRFYDVYQGEILIDHNVKDFLTGSLRNQIG 388
Qy 1372 IIPQDPVLFSGSLRMNL---DPSQYSEDEWTSLELAHLKDFVSALPDKLDHCAAGE 1428
Dh 389 LVQDDNLIFFSDYKENILGRP--DATDDEVBAKAAAHDFISLNLNGYDTEVGERGV 446
Qy 1429 NLSVGQRQLVCLARALLRKTILVLDATAVLDLTDLLIOSTIRQFEDCTVLTIAHRL 1488
Dh 447 KLSGGQQRISARIFLNNPVLILDEATSAIDLESALIIQELMDVLSKORTLTIYVHRL 506
Qy 1489 NTIMDTRVYVLDKGIQEYGAPSDLIQORGLF---YSM 1524
Dh 507 SITTHADRIYMERGRIVETGHOQLINKGAAVEHHYSI 545

RESULT 11
US-10-467-657-1346
; Sequence 1346, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASTIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657


```

: FILE REFERENCE: CL001559
: CURRENT APPLICATION NUMBER: US/10/995,561
: CURRENT FILING DATE: 2004-11-24
: NUMBER OF SEQ ID NOS: 85702
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 575
: LENGTH: 701
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-995-561-575

Query Match          4.2%  Score 419.5; DB 6; Length 701;
Best Local Similarity 25.2%  Pred. No. 4,5e-23;
Matches 149; Conservative 130; Mismatches 234; Indels 79; Gaps 21;

Qy 797 LSIPLFM--CNHVSALSNWMLSMITDDPIVINGIOEHKVRLSVYGALGISQGIANFVYS 1034
D 140 LSLFLVLVLVSSLGEMALPFFTRGRITWDILDDGSADMTTRNLTMSILITISAVLFF-- 196
Qy 1035 MAVSIG-GITLASCRLAV-----DLHSILRSPMSFFERTSGNLVNFSEKELDTVDSMI 1087
D 197 ---VGDIYNNYTMGHVSHSLQGEVFGAVLRQETEFPOQNOTGINMSRVETEDTSTLSDSL 252
Qy 1088 PEVIKMGESLFENVIGACIV-IL-----APPAAIIIPPLGLITFYFQRTVYASSRDLK 1144
D 253 SENISLFLFWYL--VRGCLLIGIMMGVSILMTVTLITLPLFLLPKVKGYQLLELYVR 310
Qy 1142 RLESVRSPPVYSHFNETLLGVSVIRAFEEORFIHOSDLKVD-----NQK---AYPSIV 1194
D 311 --EELASSQVA--IEALSMPTVRSFANEGBEAKQFEKIQEIKTLNQKAVAY---A 362
Qy 1195 ANRW-----LAVRLECVCNCIVLFAALPAVISRHSLSAGLVGLSVYSILQVTTYLAW 1246
D 363 VNSWTTISIGMLTKVGIILYIGQLVTSQA-----VSSGNLVTFVLQMGFTQAVEV 413
Qy 1247 LVKSSSEMTIVAVERLAKESERBEKAPMOIQETAPRSS---WPQVGRVFEFNKYCLRY 1302
D 414 LLSIYPRQKXVGSSEKIFEYLDRTPR-----CPSSGLLTPHLHGLVQFQDVSPAY 465
Qy 1303 --REDLPVLHINVTINGGEKVGIVGRTGAKSLSLTGLFRINBSAGEIITIDGINLXK 1360
D 466 PNRDVL-LVLOGLFTFLRPGVVTALVGENSGKSTVAAALNLQPTGGQLDGGKPLQ 524
Qy 1361 IGLHDLRFKTIIPQDPVLPFSGSLRPMN--DPFSQYSDSEEWTSLELHLKDPFYSALPDK 1418
D 525 YEHRYLHRQVAAGQEPQVFPERSIQENIAYGLTQKPMTEELTAAYVSGASHFISGLPQS 584
Qy 1419 LDHCCARGENLVSQORQLVCLARALKTKILVLDEKTAVD---LETDDLIOSTIRT 1474
D 585 YDTEVDEAGSGLSGGROAVALAPALIRKPCVLLIIDDATSALDANSQLOVQLYSES--P 642
Qy 1475 QFEDCYVLTIAHRLNTIMDYTRVIVLDGELIOEGAPESDILLQORGLFYSMAK 1526
D 643 ERYRSRVLLIQLHLSLVEQADHILFLBSGALIRBGTHOQLMEKKGCYMAWQ 694

--RESULT 14
US-10-995-561-574
: Sequence 574, Application US/10995561
: Publication No. US20050272054A1
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele et al.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
: TITLE OF INVENTION: DETECTION AND USES THEREOF
: FILE REFERENCE: CL001559
: CURRENT APPLICATION NUMBER: US/10/995,561
: CURRENT FILING DATE: 2004-11-24
: NUMBER OF SEQ ID NOS: 85702
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 574
: LENGTH: 808
: TYPE: PRT

```

[illegible]

```

Db      247 LSLFLVLVLSLGEKMLPFFTRGLTWILLQDGSADFTFRNLTLMSILTLIASAVLEF--- 303
QY      1035 MAVSIG-GILASRCLHY-----DLHSILRSPMSFFERTPSGNLVNRFSEKELDTVDSMI 1087
Db      304 ----VGDCITYNNTGWHVSHLQGEVFGAVLKRQETEFQONOTGNIMSKVETEDTSLSDSL 359
QY      1088 PEVIKMFMSLFFNVIGACIV-ILL-----ATPIAIIIPPLGLYFFVQRFYVASSRQLK 1141
Db      360 SENLFLFLMYL--VRGLCLGIMLWGSVSLTMVTLITPLFLPKVKWKYQLLEVCVR 417
QY      1142 RLESVSRSPPVSHNETLLGVSVIRAFEEQRFTHQSDLKVE---NOK---AYPSIV 1194
Db      418 --ESLAKSSOYA--IEALSAMPVRSFANEGEAKQFERKLQEIKTUNQKEAVAY---A 469
QY      1195 ANRW-----LAVRLECVNCCIIVLFAALFAVISRHSLSAGLVGLSVYSLSQVTVYLMW 1246
Db      470 VNSWTTGISGMLLVGILYIGQLVTSGA-----VSSGNLVTFVLVYOMQFTQAVEV 520
QY      1247 LVNRSSEMETNIVAVERLKEYSETEKEAPWQIQETAPPS---WPQYGRVEFRNYCLRY 1302
Db      521 LLSIYPRVQKAVGSSEKIFEXYLDRTPR-----CPSGLITPLHLEGLVQFDVSPFY 572
QY      1303 --REDLDFVLHINVTINGEKVGVGRTGAKGSLTGLFRINESAGEIITDGINIAK 1360
Db      573 PNRPDV-LVLOGLFTFLRPGEVTAIVGPNSGSKSTVALLQNLVQPTGGOLLDDGKPLPQ 631
QY      1361 IGLHDLRFKITIIPQDPVLPFSGSLRMLN--DPFSQYSDDEVWTSBELAHLKDPVSALPDK 1418
Db      632 YEHRYLHRQVAAVGOEPVFGRSLOENIAYGLTQKPTMEBITAAAVKSGAHSFISGLPQG 691
QY      1419 LDHECAGEGENTLVGQRLVCLARALLRKTKILVLDENTAAMD---LETDDLIOSTIRT 1474
Db      692 YDTEVEDEGSQLSGGQRAVALARALKRKPCVLIIDATSAIDANSQLOVEQLYES--P 749
QY      1475 QFEDCTVLTIAHRLNTIMDYTRVYVLDKGEIQEYGA PSDLLOQRGLFYSPMAK 1526
Db      750 ERYRSRYLLITQHLSLVEQADHILFLEGGAIRBEGTHOQLMEKKGCYWAMVQ 801

```

Search completed: December 15, 2005, 15:50:13
 Job time : 11.2651 secs

XX Claim 25; SEQ ID NO 6; 78pp; English.
XX The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 mutant (KR370AA) protein.
XX
SQ Sequence 1927 AA;
Query Match 100.0%; Score 9901; DB 9; Length 1927;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADSDPLMDNMTWNTSNPDFTKCFONTLVWVPCFYLMACFPFFLYLSRH 60
DB 1 MALRGFCSADSDPLMDNMTWNTSNPDFTKCFONTLVWVPCFYLMACFPFFLYLSRH 60
QY 61 DRGYIOMTPLNKTITAGFLIMYCMADLFYSFMRSGITLAPVFLVSPILLGITTLA 120
DB 61 DRGYIOMTPLNKTITAGFLIMYCMADLFYSFMRSGITLAPVFLVSPILLGITTLA 120
QY 121 TELLQLERRKGVSSGIMLTFMLVAVCALAIRSKIMTALKEBQVDFPDIFPYVES 180
DB 121 TELLQLERRKGVSSGIMLTFMLVAVCALAIRSKIMTALKEBQVDFPDIFPYVES 180
QY 181 LLLIQVLVSCFSDSPLESETIHDPNCPSSASFLSRTTFMWITGLIVRGYROPLESGD 240
DB 181 LLLIQVLVSCFSDSPLESETIHDPNCPSSASFLSRTTFMWITGLIVRGYROPLESGD 240
QY 241 LMSLNKEDTSQVVPVLYVANKKCECAKTRKQPVKVYSSKDPAQKESKTDANEVVAL 300
DB 241 LMSLNKEDTSQVVPVLYVANKKCECAKTRKQPVKVYSSKDPAQKESKTDANEVVAL 300
QY 301 IVKSPQKEMNPSLFFVLVLYKTFGRPYLMSFFPKAIHDLMMFSGPOLIKLIRKVNDRKAPD 360
DB 301 IVKSPQKEMNPSLFFVLVLYKTFGRPYLMSFFPKAIHDLMMFSGPOLIKLIRKVNDRKAPD 360
QY 361 WQGYFYTVLFEVTAQLQTLVLYHOYFHIQVSGMRITKAVIGAVYKALVITNSARKSSTV 420
DB 361 WQGYFYTVLFEVTAQLQTLVLYHOYFHIQVSGMRITKAVIGAVYKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDQRFMDLATYINMTWSAPLOVILALYLLMLNGPSVLAGVAVMLMEPVN 480
DB 421 GEIYNLMSVDQRFMDLATYINMTWSAPLOVILALYLLMLNGPSVLAGVAVMLMEPVN 480
QY 481 AVMMKTKTYOAVAHKSKDNRIKLMNELINGIKVLYAMELAFDKYLAIROBELKYLK 540
DB 481 AVMMKTKTYOAVAHKSKDNRIKLMNELINGIKVLYAMELAFDKYLAIROBELKYLK 540
QY 541 KSAVLSAVGTFWVCTPFLVALCTPAVYVTTIDENNILDAQTAFVSLAFNILRFLNLTLP 600
DB 541 KSAVLSAVGTFWVCTPFLVALCTPAVYVTTIDENNILDAQTAFVSLAFNILRFLNLTLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
QY 661 LINGITFSIPEGALVAVGVQVGGKSSLALAEADKVEGHAIVGSAVYVPOQAMIND 720
DB 661 LINGITFSIPEGALVAVGVQVGGKSSLALAEADKVEGHAIVGSAVYVPOQAMIND 720
QY 721 SLRENIILGCOLLEBYRSVIOACALPLLEILPSGDRTEIGEKVNSGGOKORVSLAR 780
DB 721 SLRENIILGCOLLEBYRSVIOACALPLLEILPSGDRTEIGEKVNSGGOKORVSLAR 780
QY 781 AVYSNADIVLPDDLPAVDAHVGKHI FENVIGPKMGLNKNKTRILVTHSMSTLPQVDVITV 840
DB 781 AVYSNADIVLPDDLPAVDAHVGKHI FENVIGPKMGLNKNKTRILVTHSMSTLPQVDVITV 840

DB 781 AVYSNADIVLPDDLPAVDAHVGKHI FENVIGPKMGLNKNKTRILVTHSMSTLPQVDVITV 840
QY 841 MSGGKISEMGSYQELIARDGAFAEFLRTVASTEOBQDAENGVTGVSQPGKEAKOMENGM 900
DB 841 MSGGKISEMGSYQELIARDGAFAEFLRTVASTEOBQDAENGVTGVSQPGKEAKOMENGM 900
QY 901 LVYDSACKQLOROLSSSSSYSGDISRRHNSTAEIOKAEAKKETWIKMEADKAOTGOVKL 960
DB 901 LVYDSACKQLOROLSSSSSYSGDISRRHNSTAEIOKAEAKKETWIKMEADKAOTGOVKL 960
QY 961 SVYWDYKKAIGLFISELSIFLFCMNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSYVG 1020
DB 961 SVYWDYKKAIGLFISELSIFLFCMNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSYVG 1020
QY 1021 ALGISQGIYAVFGYSMAVSIIGIILASRCLHYDLHSILRSPSPFEPTPSGULVNRFSKEL 1080
DB 1021 ALGISQGIYAVFGYSMAVSIIGIILASRCLHYDLHSILRSPSPFEPTPSGULVNRFSKEL 1080
QY 1081 DTVDSMIPEVIKMGSLFNVIGACIVILATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
DB 1081 DTVDSMIPEVIKMGSLFNVIGACIVILATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
QY 1141 KRLSEVSRSPTYSHFNELTIGSVYIRAEBOERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSEVSRSPTYSHFNELTIGSVYIRAEBOERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCLVLPALFAVISRHSLSAGLVGLSVSYGLQVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLCEVGNCLVLPALFAVISRHSLSAGLVGLSVSYGLQVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVRVEFRNYCLRYRDDLDFVLHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVRVEFRNYCLRYRDDLDFVLHINVTINGG 1320
QY 1321 EKVGIVRTGAGKSLTGLFRINESAGEIIDIIGINAKIGHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSLTGLFRINESAGEIIDIIGINAKIGHDLRFKTIIPQDPVLF 1380
QY 1381 SGLSRMLNDPFSQYSDSEVWTSLELAHLKDFVSAIPDLKDHCEAGGENLSVGOVOLVCL 1440
DB 1381 SGLSRMLNDPFSQYSDSEVWTSLELAHLKDFVSAIPDLKDHCEAGGENLSVGOVOLVCL 1440
QY 1441 ARALLRKTILVIDEATPAVDLETTDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ARALLRKTILVIDEATPAVDLETTDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOERYGAPSDILOORGFLPYSAKADAGLVGGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
DB 1501 DKGEIOERYGAPSDILOORGFLPYSAKADAGLVGGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
QY 1561 PRYRTERBARFVSKKGCNVAHKNIREQGRFLQDVFTTVLDLKMPTLLIFMSFLCSW 1620
DB 1561 PRYRTERBARFVSKKGCNVAHKNIREQGRFLQDVFTTVLDLKMPTLLIFMSFLCSW 1620
QY 1621 LLFAMVWMLIAFAHGDLAPGBGTNVCVTSIHSFSSAFLSIEVQVTIGFGGRWTEBCP 1680
DB 1621 LLFAMVWMLIAFAHGDLAPGBGTNVCVTSIHSFSSAFLSIEVQVTIGFGGRWTEBCP 1680
QY 1681 LAIILLYQNTVGMINAINMIGCI FMKTAQAHRAETLIPSKHAVITIRHRRLCFMRLVVG 1740
DB 1681 LAIILLYQNTVGMINAINMIGCI FMKTAQAHRAETLIPSKHAVITIRHRRLCFMRLVVG 1740
QY 1741 DLKRSMTISATIHQVVRKTTSPGEVYVPLHQVDIPMENGVGNGSIFLVAPLLIIYHVDS 1800
DB 1741 DLKRSMTISATIHQVVRKTTSPGEVYVPLHQVDIPMENGVGNGSIFLVAPLLIIYHVDS 1800
QY 1801 NSPLYDLAPSDLIHHODLEIIVILEGVEETGITTQARTSYLADEIILMGQFVPIVAED 1860
DB 1801 NSPLYDLAPSDLIHHODLEIIVILEGVEETGITTQARTSYLADEIILMGQFVPIVAED 1860
QY 1861 GRYSVDYSKFGNTIKVPPPLCTARQULDEDRSLDALTLASSRGLRAASVAVAKAKPFS 1920
DB 1861 GRYSVDYSKFGNTIKVPPPLCTARQULDEDRSLDALTLASSRGLRAASVAVAKAKPFS 1920

QY 1921 ISPDLS 1927
Db 1921 ISPDLS 1927

RESULT 2
ID ADY86938 standard; protein; 1927 AA.
AC ADY86938;
XX 02-JUN-2005 (first entry)
DE Human MRP1-mouse Klr6.2 fusion protein, SEQ ID NO: 1.
XX
XX Ionophore; biosensor; drug screening; diagnostic;
KM microorganism detection; potassium channel; fusion protein;
XX multidrug resistance protein 1; MRP1, Klr6.2.
XX Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX
XX Key Location/Qualifiers
FH Region 1..1531
FT /note= "Multidrug resistance protein 1 (MRP1)"
FT Region 1532..1537
FT /note= "Hexaglycine spacer"
FT Region 1538..1927
FT /note= "Klr6.2 protein"
XX
XX US2005063989-A1.
XX
XX 24-MAR-2005.
XX
XX 22-SEP-2003; 2003US-00665283.
XX
XX 22-SEP-2003; 2003US-00665283.
XX
XX (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX
XX Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
PI WPI; 2005-252611/26.
XX
XX New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
XX Claim 25; SEQ ID NO 1; 78bp; English.
XX
XX The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Klr6.2 protein.
XX
XX Sequence 1927 AA;

Query Match 99.9%; Score 9891; DB 9; Length 1927;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFSADGSDPLMNMNTMTNSNDPTKCFONTYLVWPCFYLMACFPFTLYSRH 60
Db 1 MALRGFSADGSDPLMNMNTMTNSNDPTKCFONTYLVWPCFYLMACFPFTLYSRH 60
QY 61 DRGYIQWTPLNKTALGFLIMIVCMADLFYFWRSGRGIPLAVFLVSPTLGITTLLA 120

Db 61 DRGYIQWTPLNKTALGFLIMIVCMADLFYFWRSGRGIPLAVFLVSPTLGITTLLA 120
QY 121 TFLIQERRRGVSSGIMLTFWVALVCAIAIIRSKITMTALKEDAQVDFRDTFYFFS 180
Db 121 TFLIQERRRGVSSGIMLTFWVALVCAIAIIRSKITMTALKEDAQVDFRDTFYFFS 180
QY 181 LLLIQVLSCPSRSLFSETHDNPCESSASFLSRITFMWTTGLIVGQRPLEGSD 240
Db 181 LLLIQVLSCPSRSLFSETHDNPCESSASFLSRITFMWTTGLIVGQRPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKQPVVYSSKDPQPKSSKVDABEVEAL 300
Db 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKQPVVYSSKDPQPKSSKVDABEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLKTFEGPYFLMSFPFKAIHIDIMFSGQIILIKFVNDTYAPD 360
Db 301 IVKSPQKEMNPSLFKVLKTFEGPYFLMSFPFKAIHIDIMFSGQIILIKFVNDTYAPD 360
QY 361 WQGYFTVLLFVTAQCLQTLVHQYFHCFSVGRIRKTAIVGAYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTAQCLQTLVHQYFHCFSVGRIRKTAIVGAYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAORFMDLATYINMWSAPLOVIALYLLMNLGSPVLAGVAVMVLAVPVN 480
Db 421 GEIVNLMSVDAORFMDLATYINMWSAPLOVIALYLLMNLGSPVLAGVAVMVLAVPVN 480
QY 481 AVNARKTKTYQVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIRQELKVLK 540
Db 481 AVNARKTKTYQVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIRQELKVLK 540
QY 541 KSAVLSAVGFTTWCCTPFLVALCTFAVYVITDENNIIADQTAFLVALFNIILPPLILP 600
Db 541 KSAVLSAVGFTTWCCTPFLVALCTFAVYVITDENNIIADQTAFLVALFNIILPPLILP 600
QY 601 MWISIVQASVSLKRLIFLSHELEBPSIERRPVKDGGGNTSYVNAATFTMARSDPT 660
Db 601 MWISIVQASVSLKRLIFLSHELEBPSIERRPVKDGGGNTSYVNAATFTMARSDPT 660
QY 661 LNCITTSIPREGALVAVVGQVCGKSSLSALALBEMDKVEGHVAIKGSVAVVPQAWIOND 720
Db 661 LNCITTSIPREGALVAVVGQVCGKSSLSALALBEMDKVEGHVAIKGSVAVVPQAWIOND 720
QY 721 SIRENIIIFGQLEBPYRSYTOACALIPDEIIPSGDRTIIGKGVNLSCGQKORVSLAR 780
Db 721 SIRENIIIFGQLEBPYRSYTOACALIPDEIIPSGDRTIIGKGVNLSCGQKORVSLAR 780
QY 781 AVYSNADIYLPDDPLASVNDHVGKHTFENVIGPKMLKNTKRIILVTHSMGYLPQVDYIIV 840
Db 781 AVYSNADIYLPDDPLASVNDHVGKHTFENVIGPKMLKNTKRIILVTHSMGYLPQVDYIIV 840
QY 841 MSGGKISEMGSYQELIARDGAPAEPLRTYASTOEODAEENGTVGSGPKKAKOMENG 900
Db 841 MSGGKISEMGSYQELIARDGAPAEPLRTYASTOEODAEENGTVGSGPKKAKOMENG 900
QY 901 LVTDGAGKQIQORLSSSSSYSDISRHNSSTAELQRAEAKKETWKLMEADKQOTGVKL 960
Db 901 LVTDGAGKQIQORLSSSSSYSDISRHNSSTAELQRAEAKKETWKLMEADKQOTGVKL 960
QY 961 SVYWDYMKAIGLIFSLFIFLFCNHYVSAIASYWYLSIMWDDPIVNGTOSHIVRSLVYG 1020
Db 961 SVYWDYMKAIGLIFSLFIFLFCNHYVSAIASYWYLSIMWDDPIVNGTOSHIVRSLVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGILASRCLVHDLHSILRSPSFERTPSGNLVNRFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGILASRCLVHDLHSILRSPSFERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMIPVYIKRMFMSIFNVIGACVILLATPIAIIIPRLGLIYFPVQRFVASSROL 1140
Db 1081 DTVDSMIPVYIKRMFMSIFNVIGACVILLATPIAIIIPRLGLIYFPVQRFVASSROL 1140
QY 1141 KRLSVSRSPVYSFNETLLGVSVIRAFEEQEFHQSOLKVDENKAYPSIVANRWLA 1200

Db 1141 KRLESVRSRSPVYSHFNETLLGVSVIRAEFEQERFIHQSDLKVDENQAKAYPSIVANRMLA 1200
 Qy 1201 VRLECVGNCLVLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETINVA 1260
 Db 1201 VRLECVGNCLVLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETINVA 1260
 Qy 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLRYRBDLDFVLRHINVTINGG 1320
 Db 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLRYRBDLDFVLRHINVTINGG 1320
 Qy 1321 EKVGIVRTGAGKSLTGLFRINESAGEIITIGINIAKIGLHDLPRKTIITIPQDPLF 1380
 Db 1321 EKVGIVRTGAGKSLTGLFRINESAGEIITIGINIAKIGLHDLPRKTIITIPQDPLF 1380
 Qy 1381 SGSLRMLNDPFSQYSDSEEWMTSLBLHDKDFVSALPDLDBHECAGGENTLVSQOROLVCL 1440
 Db 1381 SGSLRMLNDPFSQYSDSEEWMTSLBLHDKDFVSALPDLDBHECAGGENTLVSQOROLVCL 1440
 Qy 1441 ARALLRKTKIIVLDEATAVDLETFDDLIQSTIRFQEDCTVLTIAHRLNTIMDYTRVIVL 1500
 Db 1441 ARALLRKTKIIVLDEATAVDLETFDDLIQSTIRFQEDCTVLTIAHRLNTIMDYTRVIVL 1500
 Qy 1501 DKGEIOEYGAPSDLLQQRGLFYSAKADAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
 Db 1501 DKGEIOEYGAPSDLLQQRGLFYSAKADAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
 Qy 1561 PRYTRERRARFVSKGNCVNAHKNIREQGRFLQDVFTTLVDLKWPHLLIFMSFCLCSW 1620
 Db 1561 PRYTRERRARFVSKGNCVNAHKNIREQGRFLQDVFTTLVDLKWPHLLIFMSFCLCSW 1620
 Qy 1621 LLFAMWMLIAFAHGDLAGSEGTNVCVTSIHSFSSAFLEFSIEVOVTIGFGGRWTEBCP 1680
 Db 1621 LLFAMWMLIAFAHGDLAGSEGTNVCVTSIHSFSSAFLEFSIEVOVTIGFGGRWTEBCP 1680
 Qy 1681 LATILIVQNVGLMINAIMGCTFMKTAQAHRAEETLIFSKHANITLRHGLCPMLRVG 1740
 Db 1681 LATILIVQNVGLMINAIMGCTFMKTAQAHRAEETLIFSKHANITLRHGLCPMLRVG 1740
 Qy 1741 DLKSMITISATIHQOVAKTSPSEGEVVPPLHCVDPFMENGVGANGIFLVAPLIYHVIDS 1800
 Db 1741 DLKSMITISATIHQOVAKTSPSEGEVVPPLHCVDPFMENGVGANGIFLVAPLIYHVIDS 1800
 Qy 1801 NSPIYDLAPSDLHHQDLEIIVLEGVETGITTQARTSYLADEILWQGRFVPIVAED 1860
 Db 1801 NSPIYDLAPSDLHHQDLEIIVLEGVETGITTQARTSYLADEILWQGRFVPIVAED 1860
 Qy 1861 GRYSVDYKFGENTIKVPTPLCTARQDDBDRSLDMLTLASSRGPILRAASVAVAKKPKFS 1920
 Db 1861 GRYSVDYKFGENTIKVPTPLCTARQDDBDRSLDMLTLASSRGPILRAASVAVAKKPKFS 1920
 Qy 1921 ISPDLS 1927
 Db 1921 ISPDLS 1927

RESULT 3
 ADY86945
 ID ADY86945 standard; protein; 1947 AA.

XX ADY86945;
 AC 02-JUN-2005 (first entry)
 DT
 XX Human MRP1-mouse Kir6.2-HA fusion protein, SEQ ID NO: 8.
 DE
 XX
 KW Ionophore; biosensor; drug screening; diagnostic;
 KW microorganism detection; potassium channel; fusion protein;
 KW multidrug resistance protein 1; MRP1; Kir6.2.
 XX Homo sapiens.
 OS Mus musculus.
 OS Chimeric.
 OS Unidentified.

XX Key
 FH Location/Qualifiers
 FT 1.1531
 FT Region
 FT /note= "Multidrug resistance protein 1 (MRP1)"
 FT 1532.1537
 FT Region
 FT /note= "Hexaglycine spacer"
 FT 1538.1947
 FT /note= "Kir6.2 protein containing HA epitope peptide"
 FT 1651.1659
 FT /note= "HA epitope peptide"
 XX US2005063989-A1.
 XX
 XX 24-MAR-2005.
 XX
 XX 22-SEP-2003; 2003US-00665283.
 XX
 XX 22-SEP-2003; 2003US-00665283.
 XX
 XX (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 PI Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
 XX WPI; 2005-252611/26.
 XX
 XX New ion channel hybrid protein, used as electrical sensor for screening
 PT an agonist/antagonist of a membrane protein and for detecting a
 PT contaminant/pollutant in a sample.
 XX
 PS Claim 25; SEQ ID NO 8; 78bp; English.
 XX
 CC The present invention relates to a hybrid protein consisting essentially
 CC of the fusion of a membrane protein with an ion channel which is not
 CC naturally coupled to the membrane protein. The hybrid protein is used as
 CC an electrical sensor for screening of an agonist/antagonist of a membrane
 CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
 CC invention is also useful for medical diagnostics and microorganism
 CC detection. The present sequence is a fusion protein comprising human
 CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
 CC channel Kir6.2 protein containing HA epitope peptide.
 XX
 SQ Sequence 1947 AA;
 Query Match 99.6%; Score 9861; DB 9; Length 1947;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1925; Conservative 0; Mismatches 2; Indels 20; Gaps 2;
 Qy 1 MALRGFSGADGSDPLMDMNTWTNTSNPDFTKCFQNTVLLVWPCTYLLMACFPFFLYLSRH 60
 Db 1 MALRGFSGADGSDPLMDMNTWTNTSNPDFTKCFQNTVLLVWPCTYLLMACFPFFLYLSRH 60
 Qy 61 DRGYIOMTPLNKTATGFLIMIVCWADLFYSPWERSRGIFLAVPLVSPTLGTTLLA 120
 Db 61 DRGYIOMTPLNKTATGFLIMIVCWADLFYSPWERSRGIFLAVPLVSPTLGTTLLA 120
 Qy 121 TFLIQLEBRKGVQSSGIMLTFMLVALVCALALIRSKIMTALKEDAOYDLFFRDIFFYVFS 180
 Db 121 TFLIQLEBRKGVQSSGIMLTFMLVALVCALALIRSKIMTALKEDAOYDLFFRDIFFYVFS 180
 Qy 181 LLLIQVLVISCPSDRSPLEFSETHDPNCPBESSASFSLRITRWITGILVRGYROPLEBSD 240
 Db 181 LLLIQVLVISCPSDRSPLEFSETHDPNCPBESSASFSLRITRWITGILVRGYROPLEBSD 240
 Qy 241 LWSLNKEDTSROVVPVLYKMKKCKCAKTRKQPVVVVSSKQPAQKSSKVDANEVEVAL 300
 Db 241 LWSLNKEDTSROVVPVLYKMKKCKCAKTRKQPVVVVSSKQPAQKSSKVDANEVEVAL 300
 Qy 301 IVKSPQKEMNPSLFRVLYKTGPFYFLMSFFFKALHDLMMFSGPOLIKLLIFVNDTKAPD 360
 Db 301 IVKSPQKEMNPSLFRVLYKTGPFYFLMSFFFKALHDLMMFSGPOLIKLLIFVNDTKAPD 360
 Qy 361 WQGYFTYVLLFVTACLOTVLVHQYFHI CFVSGMIRIKTAVIGAVYRKALVITNSARKSTV 420

Db 361 WQGYFTVLLFVTAQCTVLVHQYFHLCFVSGMRITKTAIVAGVYRKALVITNSARKSSTV 420
 QY 421 GEIYNLMSVDAQRPMDLATYINMIWSAPLOVITLALYLMLNGSVYLAGVAVMTLMPVN 480
 Db 421 GEIYNLMSVDAQRPMDLATYINMIWSAPLOVITLALYLMLNGSVYLAGVAVMTLMPVN 480
 QY 481 AVMAKTKTYOVAMHKSNDRIKLMNEILNGI KYLKYAMELAFDPKYLAIOBELKYLK 540
 Db 481 AVMAKTKTYOVAMHKSNDRIKLMNEILNGI KYLKYAMELAFDPKYLAIOBELKYLK 540
 QY 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVITIDENNIIIDAQTAFAVSLFNILREPLNLP 600
 Db 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVITIDENNIIIDAQTAFAVSLFNILREPLNLP 600
 QY 601 MVTSSIVQASVSLKRLIFLSHELEPDSIFRRPYKOGGCTNSITTPNATTTWARSPPPT 660
 Db 601 MVTSSIVQASVSLKRLIFLSHELEPDSIFRRPYKOGGCTNSITTPNATTTWARSPPPT 660
 QY 661 LNCITFSPREGALVAVVGVCGCKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAMIOND 720
 Db 661 LNCITFSPREGALVAVVGVCGCKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAMIOND 720
 QY 721 SLRENILFGCQLEBPYRSVIOACALLPDLIELPSGDRTEIGEGVNLGGQKQVSLAR 780
 Db 721 SLRENILFGCQLEBPYRSVIOACALLPDLIELPSGDRTEIGEGVNLGGQKQVSLAR 780
 QY 781 AVYSNADITVLPDDLSAVDAVGHIFBNVIGPYGMKKNKTRILVTHSMSTLPQVDYIV 840
 Db 781 AVYSNADITVLPDDLSAVDAVGHIFBNVIGPYGMKKNKTRILVTHSMSTLPQVDYIV 840
 QY 841 MSGGKISEMGSYOELLRADGAFAEFLRTYASTEOODAEENGVTGVSGPGGEAKOMENGM 900
 Db 841 MSGGKISEMGSYOELLRADGAFAEFLRTYASTEOODAEENGVTGVSGPGGEAKOMENGM 900
 QY 901 LVTDASGKOLQROQLSSSSSYSGDISRRHNSTAEIOKAKEEFYWKLEADKAOTGOVKL 960
 Db 901 LVTDASGKOLQROQLSSSSSYSGDISRRHNSTAEIOKAKEEFYWKLEADKAOTGOVKL 960
 QY 961 SVYWDYKKAIGLFTSFLSIFLPMCNHVSAASNTWLSLMTDDPIVNGTOEHTKYRLSVYG 1020
 Db 961 SVYWDYKKAIGLFTSFLSIFLPMCNHVSAASNTWLSLMTDDPIVNGTOEHTKYRLSVYG 1020
 QY 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHYDLHSHIIRSPSPFERPSPGULVRFSEKL 1080
 Db 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHYDLHSHIIRSPSPFERPSPGULVRFSEKL 1080
 QY 1081 DTVDSMIPVYIKPMFMSLFNVIAGCIYLLATPIAIIIPPLGLIYFFVORFYVASSROL 1140
 Db 1081 DTVDSMIPVYIKPMFMSLFNVIAGCIYLLATPIAIIIPPLGLIYFFVORFYVASSROL 1140
 QY 1141 KRLSVSRSPVYSHFNETLLGVSYTRAPEOERTIHOSDLKVDENQKAYYPSIVANWMLA 1200
 Db 1141 KRLSVSRSPVYSHFNETLLGVSYTRAPEOERTIHOSDLKVDENQKAYYPSIVANWMLA 1200
 QY 1201 VRLCEVNCNIVLPALPAVISRHSLSAGVGLSVSYLOVTTYINMLVRMSSEMETIVA 1260
 Db 1201 VRLCEVNCNIVLPALPAVISRHSLSAGVGLSVSYLOVTTYINMLVRMSSEMETIVA 1260
 QY 1261 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNYCJARYREDLDFVLRIHINTINGG 1320
 Db 1261 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNYCJARYREDLDFVLRIHINTINGG 1320
 QY 1321 EKVGIVERTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
 Db 1321 EKVGIVERTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
 QY 1381 SGLSRMNLDPFSOYSDDEWVTSLELAHXDFVSALPKLHBCABEGENISVGOBOLVCL 1440
 Db 1381 SGLSRMNLDPFSOYSDDEWVTSLELAHXDFVSALPKLHBCABEGENISVGOBOLVCL 1440
 QY 1441 ARALLRKTKTILVLEDEATAVDLETDLLIOSTIRTOFEDCTVLTARHLANTIMDTRYIVL 1500
 Db 1441 ARALLRKTKTILVLEDEATAVDLETDLLIOSTIRTOFEDCTVLTARHLANTIMDTRYIVL 1500

QY 1501 DKGEIOEYGA PSDLLQORGLFYSMADKAGLVGGGGGMLSRKGIIPBEVYLTRELPDAE 1560
 Db 1501 DKGEIOEYGA PSDLLQORGLFYSMADKAGLVGGGGGMLSRKGIIPBEVYLTRELPDAE 1560
 QY 1561 PRYTRERRARFVSKKNCVNAHKNIREOGRFLQDVFPTTLVLDLKPHTLLFTMSFLCSW 1620
 Db 1561 PRYTRERRARFVSKKNCVNAHKNIREOGRFLQDVFPTTLVLDLKPHTLLFTMSFLCSW 1620
 QY 1621 LLEFAMVWMLIAFAHG-----DLAP-----GEGTNVPCVTSIHSSSAFLF 1680
 Db 1621 LLEFAMVWMLIAFAHG-----DLAP-----GEGTNVPCVTSIHSSSAFLF 1680
 QY 1661 SIEVQVITIGRGWRVTEECPLATILIVQNI VGLMIAVNLGCI FMKTAQAHRAEFLIF 1720
 Db 1661 SIEVQVITIGRGWRVTEECPLATILIVQNI VGLMIAVNLGCI FMKTAQAHRAEFLIF 1720
 QY 1721 SKHAVITLRHGRLCFMLRVGDLRKSMISATIHQVVRKTTSPGEVVPPLHQVDIPMBNG 1780
 Db 1721 SKHAVITLRHGRLCFMLRVGDLRKSMISATIHQVVRKTTSPGEVVPPLHQVDIPMBNG 1780
 QY 1781 VGGNGIFLVAPLIIVHVIDSNSPLYDLAPSDLHHODLEIIVILEGVETTGITTOARTS 1840
 Db 1801 VGGNGIFLVAPLIIVHVIDSNSPLYDLAPSDLHHODLEIIVILEGVETTGITTOARTS 1860
 QY 1841 YLADEILMGORFVPPIVAEEDGRYSVDYSKRGNTIKVPTPLCTARQLDEDSLDALTLAS 1900
 Db 1861 YLADEILMGORFVPPIVAEEDGRYSVDYSKRGNTIKVPTPLCTARQLDEDSLDALTLAS 1920
 QY 1901 SRGPLRAASVAVAKKPKFSISPSLS 1927
 Db 1921 SRGPLRAASVAVAKKPKFSISPSLS 1947
 RESULT 4
 ID ADY86941 standard; protein; 1891 AA.
 XX ADY86941;
 AC XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Human MRP1-Kir6.2 mutant delac36 fusion protein, SEQ ID NO: 4.
 XX
 KW Ionophore; biosensor; drug screening; diagnostic;
 KW microorganism detection; potassium channel; fusion protein;
 KW multidrug resistance protein 1; MRP1; Kir6.2; mutein.
 OS Homo sapiens.
 OS Mus musculus.
 OS Chimeric.
 XX
 XX Key
 FT Region 1..1531
 FT /note="Multidrug resistance protein 1 (MRP1)"
 FT Region 1532..1537
 FT /note="Hexaglycine spacer"
 FT Region 1538..1891
 FT /note="Kir6.2 mutant delac36 protein"
 XX
 PN US2005063989-A1.
 XX
 PD 24-MAR-2005.
 XX
 PF 22-SEP-2003; 2003US-00665283.
 XX
 PR 22-SEP-2003; 2003US-00665283.
 XX
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 XX
 PI Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
 XX WPI; 2005-252611/26.

XX New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
PS Claim 25; SEQ ID NO 4; 78bp; English.
XX
CC The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 mutant detac36 protein.
XX
SQ Sequence 1891 AA.
Query Match 98.3%; Score 9734; DB 9; Length 1891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALGFCSDGSDSLPMDMNTWNTSNPDTKCFQNTVLWVPCFLYMACFPFELYLSRH 60
DB 1 MALGFCSDGSDSLPMDMNTWNTSNPDTKCFQNTVLWVPCFLYMACFPFELYLSRH 60
QY 61 DRGIQMTPLNKKTALGFLIMIVCMADLFYSFMRSGIFLAPFLVSPILGITTLA 120
DB 61 DRGIQMTPLNKKTALGFLIMIVCMADLFYSFMRSGIFLAPFLVSPILGITTLA 120
QY 121 TFLIQLEBRKGVSGGIMLTFMLVALYCALALRSKIMTALKEDAQVDLPFDITFYYS 180
DB 121 TFLIQLEBRKGVSGGIMLTFMLVALYCALALRSKIMTALKEDAQVDLPFDITFYYS 180
QY 181 LLLIQVLVSCSDSPLESETIHDNPPCRESSASLSTITFWITGLIVRGYRQPLESD 240
DB 181 LLLIQVLVSCSDSPLESETIHDNPPCRESSASLSTITFWITGLIVRGYRQPLESD 240
QY 181 LLLIQVLVSCSDSPLESETIHDNPPCRESSASLSTITFWITGLIVRGYRQPLESD 240
DB 181 LLLIQVLVSCSDSPLESETIHDNPPCRESSASLSTITFWITGLIVRGYRQPLESD 240
QY 241 LMSLNKEDTSBOQVPLVAKMKKECAKTRKQPVKVVYSSKPAOPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSBOQVPLVAKMKKECAKTRKQPVKVVYSSKPAOPKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLVYKTFGPFYFLMSFPFKAIDHLMFSGPQLIKLIRVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFKVLVYKTFGPFYFLMSFPFKAIDHLMFSGPQLIKLIRVNDTKAPD 360
QY 361 WQGFYTYVLFTVYACLOTLVHOYFHICFVSGMRKTKAVIGAVYRKALVITNSARKSSTV 420
DB 361 WQGFYTYVLFTVYACLOTLVHOYFHICFVSGMRKTKAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAORFMDLATYINMISAPLOYILALYLMLNGPSVLGAVVMVLMPVN 480
DB 421 GEIVNLMSVDAORFMDLATYINMISAPLOYILALYLMLNGPSVLGAVVMVLMPVN 480
QY 481 AYMAKTKTYOVANHKSKDNRIKLANEILNGIKVLKLYAMELAFKQVLAIROBELKYLK 540
DB 481 AYMAKTKTYOVANHKSKDNRIKLANEILNGIKVLKLYAMELAFKQVLAIROBELKYLK 540
QY 541 KSAIYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFLVSLAFNLIKPLNLTLP 600
DB 541 KSAIYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFLVSLAFNLIKPLNLTLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGNSTIVRAATFTWASDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGNSTIVRAATFTWASDPT 660
QY 661 LAGTTFSPREGALVAVVGQVCGKSSLLSALLAEMDKYEGHVAIKGSVAYVPOQMIQND 720
DB 661 LAGTTFSPREGALVAVVGQVCGKSSLLSALLAEMDKYEGHVAIKGSVAYVPOQMIQND 720
QY 721 SLRENIILFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQKORVSLAR 780
DB 721 SLRENIILFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQKORVSLAR 780

DB 721 SLRENIILFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQKORVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDAAVGHKIPENYIGPKMKLNKTRILVTHSMSTLPQVDVIY 840
DB 781 AVYSNADIYLFDDPLSAVDAAVGHKIPENYIGPKMKLNKTRILVTHSMSTLPQVDVIY 840
QY 841 MSGGKISEMGSYQELARDGAFAEFLTYASTBOEDAENGVTGVSQPKAKOMENG 900
DB 841 MSGGKISEMGSYQELARDGAFAEFLTYASTBOEDAENGVTGVSQPKAKOMENG 900
QY 901 LVYDSAKCOIQRQSSSSSYSGDISRHNSSTAELQKAEKKEETWKMEADKATGVYK 960
DB 901 LVYDSAKCOIQRQSSSSSYSGDISRHNSSTAELQKAEKKEETWKMEADKATGVYK 960
QY 961 SVYDWYKAKIGLFISSIFLFMCNHSALASNYMSTLMTDDPIVNGTOEHTKRLSYVG 1020
DB 961 SVYDWYKAKIGLFISSIFLFMCNHSALASNYMSTLMTDDPIVNGTOEHTKRLSYVG 1020
QY 1021 ALGISGIIAVFGYSMAVSIQGIILASRCLAYDLHSIIRSPMSFPERTPSGNLVNRFKEL 1080
DB 1021 ALGISGIIAVFGYSMAVSIQGIILASRCLAYDLHSIIRSPMSFPERTPSGNLVNRFKEL 1080
QY 1081 DTVDSMTPEVIXMPMGSLFNVIGACIVILATPIAIIIPPLGLIYFFVQRFYASSROL 1140
DB 1081 DTVDSMTPEVIXMPMGSLFNVIGACIVILATPIAIIIPPLGLIYFFVQRFYASSROL 1140
QY 1141 KRLESVSRSPYSHFNETLIGSVYIRAFEBEOREIHOSDKVDENOKAYPSIVANFMA 1200
DB 1141 KRLESVSRSPYSHFNETLIGSVYIRAFEBEOREIHOSDKVDENOKAYPSIVANFMA 1200
QY 1201 VRLFCVGNCTYLPAALPAVYSRHSLSAGLVGSYSLOVTTYANMLVRSSSEMETNIVA 1260
DB 1201 VRLFCVGNCTYLPAALPAVYSRHSLSAGLVGSYSLOVTTYANMLVRSSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVEFRNYCLRYREDLPVLVRLHINVTNGG 1320
DB 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVEFRNYCLRYREDLPVLVRLHINVTNGG 1320
QY 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRNMLDPEFSQYSDSEVWTSLELAHKDFVSALPKLHCEKAGEGENTSVGQROLVCL 1440
DB 1381 SGSLRNMLDPEFSQYSDSEVWTSLELAHKDFVSALPKLHCEKAGEGENTSVGQROLVCL 1440
QY 1441 ARALLRKTILVDEATAVADLETDDLIOSTIRQFEDCTVLTIAHRLNTIMDTRYVL 1500
DB 1441 ARALLRKTILVDEATAVADLETDDLIOSTIRQFEDCTVLTIAHRLNTIMDTRYVL 1500
QY 1501 DKGEIOEYGAPSDLLQGRGLFYSMAKDAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
DB 1501 DKGEIOEYGAPSDLLQGRGLFYSMAKDAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
QY 1561 PRYTRERRARFVSKKNCNVAHKNIREQGFLODVFTTIVDLKWPHTLLIFTSFLCSW 1620
DB 1561 PRYTRERRARFVSKKNCNVAHKNIREQGFLODVFTTIVDLKWPHTLLIFTSFLCSW 1620
QY 1621 LLFAMVWMLIAFAHGDLAPBGNTVPCVTSIHSFSSAPLFSIEVQVTIGFGGRVATECP 1680
DB 1621 LLFAMVWMLIAFAHGDLAPBGNTVPCVTSIHSFSSAPLFSIEVQVTIGFGGRVATECP 1680
QY 1681 LAIILIVONVGMINAINMIGCFEMKTAQAHRAEFLIRSKAVITLRBGRLCFMLERVG 1740
DB 1681 LAIILIVONVGMINAINMIGCFEMKTAQAHRAEFLIRSKAVITLRBGRLCFMLERVG 1740
QY 1741 DLKRSMTISATIHQVVRKTSPEGEVPLHQVDIPENGVGNGIFLVAELIITYHVIDS 1800
DB 1741 DLKRSMTISATIHQVVRKTSPEGEVPLHQVDIPENGVGNGIFLVAELIITYHVIDS 1800
QY 1801 NSPLYDLAPSDLHHODELIIIVILEGVETTGITTOARTSYLADEIILMGQRFVPIVAED 1860
DB 1801 NSPLYDLAPSDLHHODELIIIVILEGVETTGITTOARTSYLADEIILMGQRFVPIVAED 1860

QY 1861 GRYSVDYSKFGNTIKVFTPLCTARQLDEDRS 1891
 DB 1861 GRYSVDYSKFGNTIKVFTPLCTARQLDEDRS 1891

RESULT 5
 ID AAM57486 standard; protein; 1531 AA.
 AAM57486
 AC AAM57486;
 DT 14-AUG-1998 (first entry)
 DE Human MRP variant 1cPgpa (Lei/Pgpa).
 KM Multidrug resistance-associated protein; MRP; tumour; human; variant;
 XX multidrug resistance; MDR; leishmania P-glycoprotein; 1cPgpa; Lei/Pgpa.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 685 /label= I685S
 FT /note= "wild-type Leu is replaced by Ser"
 FT Misc-difference 1282 /label= R1282A
 FT /note= "wild-type Arg is replaced by Ala"
 XX US5766880-A.
 PD 16-JUN-1998.
 XX 05-JUN-1995; 95US-00463092.
 XX 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX (TOOH) UNIV QUEBENS KINGSTON.
 PA Cole SP, Deeley RG;
 PI WPI; 1998-361687/31.
 DR N-PSDB; AAV31498.
 XX DNA encoding protein associated with multi-drug resistance - useful for
 PT as probe for identifying multi-drug resistant tumour cells.
 PS Claim 1; Col 67-78; 82pp; English.
 XX This represents a variant of the human multidrug resistance-associated
 CC protein (MRP). This natural variant is a leishmania P-glycoprotein related
 CC molecule 1cPgpa (Lei/Pgpa). The human and murine MRP nucleic acid
 CC molecules can be used as probes for identifying multidrug resistant
 CC tumour cells by hybridisation to mRNA from tumour cells. The antisense
 CC nucleic acid can be used to reverse multidrug resistance (MDR). A
 CC recombinant expression vector containing the MRP nucleic acid molecules
 CC operatively linked to at least one regulatory sequence can be used to
 CC transform a host cell to produce a recombinant MDR-associated protein
 XX Sequence 1531 AA;
 SQ

Query Match 79.4%; Score 7860; DB 2; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARGPESADGSPIDMNTNTNTSNDFTKCFONTYLVVWPCFYLMACFPFFLYLSRH 60
 DB 1 MARGPESADGSPIDMNTNTNTSNDFTKCFONTYLVVWPCFYLMACFPFFLYLSRH 60
 QY 61 DRGYIQWTPLNKTALGFLMTIVCWADLFYSFWEBSRGIFLAPVFLVSPTLTGITTLA 120

DB 61 DRGYIQWTPLNKTALGFLMTIVCWADLFYSFWEBSRGIFLAPVFLVSPTLTGITTLA 120
 QY 121 TPLIQLERRRGVSSGIMLTFWLVALVCAIAIRSKITMTLKEDAQVDLFRDITFFYYFS 180
 DB 121 TPLIQLERRRGVSSGIMLTFWLVALVCAIAIRSKITMTLKEDAQVDLFRDITFFYYFS 180
 QY 181 LLLIQLVLSGFSRSPLEFSTTHDPNCPRESSASFLSRITFWMTGLIVNGYQPLGSD 240
 DB 181 LLLIQLVLSGFSRSPLEFSTTHDPNCPRESSASFLSRITFWMTGLIVNGYQPLGSD 240
 QY 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKOPVYVSSKDPAPQPKSSKYDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKOPVYVSSKDPAPQPKSSKYDANEVEAL 300
 QY 301 IVKSPQKEMNPSPFKVLYKTFGPFYFMSPFKAIHIDIMFSGPQLTKLTKFVNDTAPD 360
 DB 301 IVKSPQKEMNPSPFKVLYKTFGPFYFMSPFKAIHIDIMFSGPQLTKLTKFVNDTAPD 360
 QY 361 WOGYFTYVLLFVTRACLOTVLHQYFHCYFSGKRIKTAIVGAYRKALVITNSARKSTV 420
 DB 361 WOGYFTYVLLFVTRACLOTVLHQYFHCYFSGKRIKTAIVGAYRKALVITNSARKSTV 420
 QY 421 GEIYNLMSVDAORFMDLATYINNIWSAPLOVILALYLMNLGPSVLAGVAVMVLAMPVN 480
 DB 421 GEIYNLMSVDAORFMDLATYINNIWSAPLOVILALYLMNLGPSVLAGVAVMVLAMPVN 480
 QY 481 AVNARKTKTYQVAMKSKDNRIKLMEILNGIKVLKLYAMELAFKDVLAIRBELVVK 540
 DB 481 AVNARKTKTYQVAMKSKDNRIKLMEILNGIKVLKLYAMELAFKDVLAIRBELVVK 540
 QY 541 KSAVLSAVGTFPTWCPTPLVALCTFAVYVITDENNIIADQTAVSALFNILRPLNLP 600
 DB 541 KSAVLSAVGTFPTWCPTPLVALCTFAVYVITDENNIIADQTAVSALFNILRPLNLP 600
 QY 601 MVISSIVQASVSLKRLIFLSHELEPDSIERPVDGGGTNSITVYNAFTFARSPPT 660
 DB 601 MVISSIVQASVSLKRLIFLSHELEPDSIERPVDGGGTNSITVYNAFTFARSPPT 660
 QY 661 LMGITFSIPREGALVAVVGQVCGCKSSLSLALMEKVEBVAIKGSVAVYPOAMVQND 720
 DB 661 LMGITFSIPREGALVAVVGQVCGCKSSLSLALMEKVEBVAIKGSVAVYPOAMVQND 720
 QY 721 SIRENIIFGCQLEPEPYRSVYQACALLPDLEIIPSGDRTFIEGKGVNLSCGQKORVSLR 780
 DB 721 SIRENIIFGCQLEPEPYRSVYQACALLPDLEIIPSGDRTFIEGKGVNLSCGQKORVSLR 780
 QY 781 AVYSNADITYLPDPLSAVDHVGKHPENVIGPKMKLNKTRILVTHSMGYLPOVDVIV 840
 DB 781 AVYSNADITYLPDPLSAVDHVGKHPENVIGPKMKLNKTRILVTHSMGYLPOVDVIV 840
 QY 841 MSGGKISEMGSYOELLARDGAFAPLRTYASTEOBDAENGTVGSGPKEAKOMENG 900
 DB 841 MSGGKISEMGSYOELLARDGAFAPLRTYASTEOBDAENGTVGSGPKEAKOMENG 900
 QY 901 LVTDSAGKOLQROLSSSSSYSGDISRHNSSTAELQAEAKKETWKLMEADKAQTCQVKL 960
 DB 901 LVTDSAGKOLQROLSSSSSYSGDISRHNSSTAELQAEAKKETWKLMEADKAQTCQVKL 960
 QY 961 SYVWDYMKALGLFISPLSTFLFMCNIVSALASNYWLSIMTDDPIVNGTOHTVRLSVYG 1020
 DB 961 SYVWDYMKALGLFISPLSTFLFMCNIVSALASNYWLSIMTDDPIVNGTOHTVRLSVYG 1020
 QY 1021 ALGISOGIAVFGYSMAVSIIGILASRCLVLDLHSLRSPMSFERTPSGNLVNRSKEL 1080
 DB 1021 ALGISOGIAVFGYSMAVSIIGILASRCLVLDLHSLRSPMSFERTPSGNLVNRSKEL 1080
 QY 1081 DTVDSMIPEVIKMFMSLBNVIGACIVILLATPIAIIIPPLGLIYFVQRFVASSROL 1140
 DB 1081 DTVDSMIPEVIKMFMSLBNVIGACIVILLATPIAIIIPPLGLIYFVQRFVASSROL 1140
 QY 1141 KRLSEYSRSPVSHFNETLLGVSVIRAFEOERFHQSGLKVDENOKAYPSIVARWLA 1200

Db 1141 KRLESVRSVPYSHFNETLGVSVIRAEFEQERFIHOSDLKVDENQAYPSIVANRMLA 1200
Qy 1201 VRLCEVGNCTVLFPAALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRSSEMETNIVA 1260
Db 1201 VRLCEVGNCTVLFPAALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPWQIQETAPPSWQVGRVEFRNYCLARYEDDLDFVLHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIQETAPPSWQVGRVEFRNYCLARYEDDLDFVLHINVTINGG 1320
Qy 1321 EKVGIVERTGAGKSLTGLFRINESANGETIIDGINIAKIGLHDLARKITIIIPDPLTF 1380
Db 1321 EKVGIVERTGAGKSLTGLFRINESANGETIIDGINIAKIGLHDLARKITIIIPDPLTF 1380
Qy 1381 SGLSRMNLDPFSQYSDSEEWMTSLBLAKDFVSALPDKLDEHCEAGGENSEVGQROLVCL 1440
Db 1381 SGLSRMNLDPFSQYSDSEEWMTSLBLAKDFVSALPDKLDEHCEAGGENSEVGQROLVCL 1440
Qy 1441 APALIRKTKIIVDBATPAVDLETDDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRIVL 1500
Db 1441 APALIRKTKIIVDBATPAVDLETDDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRIVL 1500
Qy 1501 DKGEIOEYGAAPSDLLQQRGLFYMAKQAGLV 1531
Db 1501 DKGEIOEYGAAPSDLLQQRGLFYMAKQAGLV 1531

RESULT 6

AAW74471 standard; protein, 1531 AA.

AAW74471;

18-MAY-1999 (first entry)

Human multidrug resistance-associated protein variant.

Multidrug resistance-associated protein; MDR; human; diagnosis;

MDR tumour cell identification; cancer therapy.

Homo sapiens.

US5882875-A.

16-MAR-1999.

05-JUN-1995; 95US-00462109.

27-OCT-1992; 92US-00966923.

08-MAR-1993; 93US-00029340.

26-OCT-1993; 93US-00141893.

20-MAR-1995; 95US-00407207.

(TOOH) UNIV QUEBENS KINGSTON.

Cole SPC, Deeley RG;

WPI: 1999-214061/18.

N-PSDB; AAX21977.

Identifying a multidrug resistant tumor cell by contacting the cell with

an antibody/antigen-binding fragment - which binds to an expressed

protein encoded by multidrug resistance-associated protein (MRP) nucleic

acid.

Claim 3; Col 69-80; 80pp; English.

CC for inhibiting/killing a MDR tumour cell can be useful for treating
CC breast cancer, leukemias, fibrosarcomas, cervical cancer, gliomas,
CC thymomas, neuroblastomas and lung cancer. The MDR DNA sequences when
CC labeled are useful as molecular probes for diagnosing multidrug
CC resistance of a tumour (using cells from a tumour biopsy) and for
CC designing ribozymes which are capable of cleaving a single-stranded
CC nucleic acid encoding a protein having MRP activity. Recombinant
CC expression vectors containing human MDR coding sequences can be
CC transfected into a drug sensitive cell line to produce a protein in the
CC cell which confers MDR, protecting non-resistant non-tumour cells from
CC the effects of chemotherapeutics has major clinical importance. Cells
CC transformed with the MDR coding sequences are useful for testing
CC potential therapeutic agents for their effectiveness against MDR cells
CC and for identifying chemosensitizers of a therapeutic agent
XX
SQ Sequence 1531 AA:
Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRGFCSADGSDPLMDMNTVMTNSNDPTKCFQNTVLVWPVPCFYLWACPFYFLYSRH 60
Db 1 MALRGFCSADGSDPLMDMNTVMTNSNDPTKCFQNTVLVWPVPCFYLWACPFYFLYSRH 60
Qy 61 DRGYIQTPLNKTKTALGFLIMVCMADLFYSFWERSRGIFLADVFLVSPPTLIGITTLA 120
Db 61 DRGYIQTPLNKTKTALGFLIMVCMADLFYSFWERSRGIFLADVFLVSPPTLIGITTLA 120
Qy 121 TFLIQLEBRKGVSSGIMLTFMLVALYCALAILSKIMTAKENAVDLPDIDIFYFYFS 180
Db 121 TFLIQLEBRKGVSSGIMLTFMLVALYCALAILSKIMTAKENAVDLPDIDIFYFYFS 180
Qy 181 LLLIQVLVSCPSDSSPLFSETIHDNCPBESSAFLSRIFPMWTGLIVRGYRPLEGSD 240
Db 181 LLLIQVLVSCPSDSSPLFSETIHDNCPBESSAFLSRITFTWITGLIVRGYRPLEGSD 240
Qy 241 LMSLNKEDTSQVVPVIVKWKKECAKTRKQPVKVYSSKQPAQPKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVIVKWKKECAKTRKQPVKVYSSKQPAQPKSSKYDANEVEAL 300
Qy 301 IVKSPQKEMNPSLKVLYKTGPFLMSFPFKAIHDLMSRGPOLIKLLIFVNDTRAPD 360
Db 301 IVKSPQKEMNPSLKVLYKTGPFLMSFPFKAIHDLMSRGPOLIKLLIFVNDTRAPD 360
Qy 361 WQGFYTYLAFVTAQLOTLVLAHQYFHLCEVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGFYTYLAFVTAQLOTLVLAHQYFHLCEVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLSVDAQRPMDLATYINMIMSAPLQVILALYLLMLNLAGSVLAGVAVMYLMPVN 480
Db 421 GEIVNLSVDAQRPMDLATYINMIMSAPLQVILALYLLMLNLAGSVLAGVAVMYLMPVN 480
Qy 481 AVAMAKTKTYOVAAHKSNDRIKLMNEILNGIKYLKUYAMELAKDKYLAIOBELKYLK 540
Db 481 AVAMAKTKTYOVAAHKSNDRIKLMNEILNGIKYLKUYAMELAKDKYLAIOBELKYLK 540
Qy 541 KSAVLSAAGTFTWCTPFLVALCTPAVYVITDENNNILDAOTAPVSLAFNLRPLNTLP 600
Db 541 KSAVLSAAGTFTWCTPFLVALCTPAVYVITDENNNILDAOTAPVSLAFNLRPLNTLP 600
Qy 601 MVISIVQASVSLRLRIFLSHEELEPDSIRRPVKDGGTNSITVRNATFTWASDPT 660
Db 601 MVISIVQASVSLRLRIFLSHEELEPDSIRRPVKDGGTNSITVRNATFTWASDPT 660
Qy 661 LNGITFSPREGALVAVVQVCGKSSLLSALLAEMDVKEGHVAIKGSVAVYPQAMIOND 720
Db 661 LNGITFSPREGALVAVVQVCGKSSLLSALLAEMDVKEGHVAIKGSVAVYPQAMIOND 720
Qy 721 SLRENIILFGCLLEPPYRSVQAQALLPDLEILSFGPTEIGEKGUNVSGGOKORVSLAR 780
Db 721 SLRENIILFGCLLEPPYRSVQAQALLPDLEILSFGPTEIGEKGUNVSGGOKORVSLAR 780

QY 781 AVYSNADIIYLPDDLSAVDAHVGKHIENIVGPKGMLKNKTRILVTHSMSTLPOVDVIV 840
 DB 781 AVYSNADIIYLPDDLSAVDAHVGKHIENIVGPKGMLKNKTRILVTHSMSTLPOVDVIV 840
 QY 841 MSGKIKEMSGSYQGLIARDGAFAPFLTYASTBODAEENGVTGSGPGKEAQMENGM 900
 DB 841 MSGKIKEMSGSYQGLIARDGAFAPFLTYASTBODAEENGVTGSGPGKEAQMENGM 900
 QY 901 LVTDSAKOIQOROLSSSSSSSGDISRHHNSTAEIOKAKEEFMKMEADKATQGVKL 960
 DB 901 LVTDSAKOIQOROLSSSSSSSGDISRHHNSTAEIOKAKEEFMKMEADKATQGVKL 960
 QY 961 SYVDYMKAIQGLFISFLSIFLMCNHVSALASNYWLSMTDPIVNGTOEHTKRLSYVG 1020
 DB 961 SYVDYMKAIQGLFISFLSIFLMCNHVSALASNYWLSMTDPIVNGTOEHTKRLSYVG 1020
 QY 1021 ALGISGIIAVFGSMAVSIIGIILASRCLHVDLHSLIRSPMSFEKTPSGNLVNRFSKEL 1080
 DB 1021 ALGISGIIAVFGSMAVSIIGIILASRCLHVDLHSLIRSPMSFEKTPSGNLVNRFSKEL 1080
 QY 1081 DTVDMSIPEVYKMFMSLFFNYGACITILATPIAIIIPPLGLIYFFVQAFYVASSROL 1140
 DB 1081 DTVDMSIPEVYKMFMSLFFNYGACITILATPIAIIIPPLGLIYFFVQAFYVASSROL 1140
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYYSIVANWMLA 1200
 DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYYSIVANWMLA 1200
 QY 1201 VRLECVCNCIYLFALPAVIRSRHSLASGLVGLSVYSLSQVTTYLNLVMSSEMETNIVA 1260
 DB 1201 VRLECVCNCIYLFALPAVIRSRHSLASGLVGLSVYSLSQVTTYLNLVMSSEMETNIVA 1260
 QY 1261 VERIKESVETKEKAPMOIOETAPSSWPQVGRVFRNYCLARYEDLPVLRHINVTINGG 1320
 DB 1261 VERIKESVETKEKAPMOIOETAPSSWPQVGRVFRNYCLARYEDLPVLRHINVTINGG 1320
 QY 1321 EKVGIIVRTGAGKSSLLTGLFRINESAGEIIDIINIAKIGLHDLFPKTIIPQDVLFF 1380
 DB 1321 EKVGIIVRTGAGKSSLLTGLFRINESAGEIIDIINIAKIGLHDLFPKTIIPQDVLFF 1380
 QY 1381 SGLSRMLDPPSOYSDEEVTSLLELAHKDFVSALPKLDHECAEGENLSVGRQVLCL 1440
 DB 1381 SGLSRMLDPPSOYSDEEVTSLLELAHKDFVSALPKLDHECAEGENLSVGRQVLCL 1440
 QY 1441 ARAALRRTKILVDEAFAVADLETDDLIQSTIRTOFEDCTVLTAHRLANTIMDYTRIVL 1500
 DB 1441 ARAALRRTKILVDEAFAVADLETDDLIQSTIRTOFEDCTVLTAHRLANTIMDYTRIVL 1500
 QY 1501 DKGEIOEYGAPSDLLQORGLFYSAKXDAAGLV 1531
 DB 1501 DKGEIOEYGAPSDLLQORGLFYSAKXDAAGLV 1531

XX 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX (TOOH) UNIV QUEBENS KINGSTON.
 PI Cole SPC, Dealey RG;
 XX WPI; 1999-253868/21.
 DR N-PSDB; AAX19818.
 PT Protecting mammalian cells against cytotoxic drugs.
 XX Claim 2; Col 79-86; 82pp; English.
 CC The present sequence represents a human multidrug resistance-associated
 CC protein (MRP). The present invention also describes a method for
 CC protecting a mammalian cell against the cytotoxicity of antineoplastic,
 CC epipodophylotoxins and Vinca alkaloids (A) by introducing into it a
 CC nucleic acid (I) that hybridizes under stringent conditions to a nucleic
 CC acid (II) that encodes an MRP protein (III). Introduction of (I) protects
 CC cells against cytotoxic effects of (A), particularly to protect normal
 CC cells against (A) being used for treatment of cancers. Cells transformed
 CC with (I) can be used to screen for agents that affect multidrug
 CC resistance or are directly toxic to multidrug resistant cells, i.e.
 CC potential therapeutics for multidrug-resistant cancers. Confering
 CC resistance to normal cells should allow an increase in the dose of (A)
 CC that can be administered safely
 CC XX
 SQ Sequence 1531 AA:
 Query Match 79.4%; Score 7860; DB 2; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MALRGFCSAGSDPMDMNTNTNTPNDFPKTCQNTYLVVWPCFYLMACPPFYLYSRH 60
 DB 1 MALRGFCSAGSDPMDMNTNTNTPNDFPKTCQNTYLVVWPCFYLMACPPFYLYSRH 60
 QY 61 DRGIQMTPLNKTKITLGLFLLWIVCMADLFYSFMEBSRGFLAPVPLVSPTLGITLLA 120
 DB 61 DRGIQMTPLNKTKITLGLFLLWIVCMADLFYSFMEBSRGFLAPVPLVSPTLGITLLA 120
 QY 121 TFLIQERRKGVSSGIMLTFWLVALVICALAIIRSKIMTALKEDAQVDFRDTFFYYFS 180
 DB 121 TFLIQERRKGVSSGIMLTFWLVALVICALAIIRSKIMTALKEDAQVDFRDTFFYYFS 180
 QY 181 LLLIQVLSCFSRSPFLFSTIHDNPNCPSSASFLSRITFWWITGLIYNGRQPLEGSD 240
 DB 181 LLLIQVLSCFSRSPFLFSTIHDNPNCPSSASFLSRITFWWITGLIYNGRQPLEGSD 240
 QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRQPKVYSSDPPQPPRESSVVDNNEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRQPKVYSSDPPQPPRESSVVDNNEVEAL 300
 QY 301 IVKSPQKWNPSLFLKVLTKTFCGYPLMSFFFKAIHDLMPSCQIILKLIFVNDTKAPD 360
 DB 301 IVKSPQKWNPSLFLKVLTKTFCGYPLMSFFFKAIHDLMPSCQIILKLIFVNDTKAPD 360
 QY 361 WQGYFTVLLFVYACIQTLYLHOYFHIQVSGMRITAVIGAYRRALVITNSARKSSTV 420
 DB 361 WQGYFTVLLFVYACIQTLYLHOYFHIQVSGMRITAVIGAYRRALVITNSARKSSTV 420
 QY 421 GEIVNLMSVDAORFMDLATYINNIWSAPLOVITATLVMLNLPSPVLAGAVVWLVNVPV 480
 DB 421 GEIVNLMSVDAORFMDLATYINNIWSAPLOVITATLVMLNLPSPVLAGAVVWLVNVPV 480
 QY 481 AVAMKTKTYQVAHMSKONRIKLANEILNGIKVTLKYAMELAFKDKVLAIROBELKVLK 540
 DB 481 AVAMKTKTYQVAHMSKONRIKLANEILNGIKVTLKYAMELAFKDKVLAIROBELKVLK 540

```

QY 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVITDENNIIDAQTAFAVSLALFNILRPELILP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVITDENNIIDAQTAFAVSLALFNILRPELILP 600
QY 601 MVTISITVOASVSLRLRPLFLSHELEPDSIRERRPVKGGGNSITVNAATWTARSDPT 660
DB 601 MVTISITVOASVSLRLRPLFLSHELEPDSIRERRPVKGGGNSITVNAATWTARSDPT 660
QY 661 LNCITFSP1PEGALVAVVGVQVCGKSSLLSALLAEKDKEGHVAIKGSAAVYPQQAQND 720
DB 661 LNCITFSP1PEGALVAVVGVQVCGKSSLLSALLAEKDKEGHVAIKGSAAVYPQQAQND 720
QY 721 SLRENIIFGCQLEBPYRSVIQACALLPDLLEIPSGDRTIEGKGVNLSGGQKORVSLAR 780
DB 721 SLRENIIFGCQLEBPYRSVIQACALLPDLLEIPSGDRTIEGKGVNLSGGQKORVSLAR 780
QY 781 AAVSNADIYLPDDPLSAVDAAVGHKIFENVIQPGKMLKNKTRILVTHSMSTLPQVDVITV 840
DB 781 AAVSNADIYLPDDPLSAVDAAVGHKIFENVIQPGKMLKNKTRILVTHSMSTLPQVDVITV 840
QY 841 MSGGKISEMGSYQELIARDGAFAEFLRTYASTEOQDAEENGVTGVSQPGKEAKOMENGM 900
DB 841 MSGGKISEMGSYQELIARDGAFAEFLRTYASTEOQDAEENGVTGVSQPGKEAKOMENGM 900
QY 901 LVTDAGKQIQRQLSSSSSSSGDISRHHNSTAEIQAQAKKEETWKLMEADKAOTGYKL 960
DB 901 LVTDAGKQIQRQLSSSSSSSGDISRHHNSTAEIQAQAKKEETWKLMEADKAOTGYKL 960
QY 961 SVYWDYKAIQGLFTSFLSIFLMCNHVALASNWLSTWTDPIVNGQHEHTKRLSYG 1020
DB 961 SVYWDYKAIQGLFTSFLSIFLMCNHVALASNWLSTWTDPIVNGQHEHTKRLSYG 1020
QY 1021 ALGISQGIAVFGYMAVSIQGIILASRCLHVDLHLSILSPSFERTPSGMLVNRFSKEL 1080
DB 1021 ALGISQGIAVFGYMAVSIQGIILASRCLHVDLHLSILSPSFERTPSGMLVNRFSKEL 1080
QY 1081 DTVDSMIPEVIKMGMSLFENVIGACTIVILLATPIAIIIPGLGIYFFVQSPFYVASSRQL 1140
DB 1081 DTVDSMIPEVIKMGMSLFENVIGACTIVILLATPIAIIIPGLGIYFFVQSPFYVASSRQL 1140
QY 1141 KRLESVRSRPSYSHENETLGVSVIRAFEBEERFTHQDLDVDEKQAYYSIVANRWLA 1200
DB 1141 KRLESVRSRPSYSHENETLGVSVIRAFEBEERFTHQDLDVDEKQAYYSIVANRWLA 1200
QY 1201 VRLECVGNCIVLPALFAVISRHSISAGLVSVSYSQVTTYLNLVVRMSEMETNIVA 1260
DB 1201 VRLECVGNCIVLPALFAVISRHSISAGLVSVSYSQVTTYLNLVVRMSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIQETAPPSWPQVRVEFRNYCLARYREDDLPFLRHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIQETAPPSWPQVRVEFRNYCLARYREDDLPFLRHINVTINGG 1320
QY 1321 EKVGIVERTGAGKSLTIGLFRINESAGEIIIDGINIAKIGLHDLRKTITIIIPDDPLTF 1380
DB 1321 EKVGIVERTGAGKSLTIGLFRINESAGEIIIDGINIAKIGLHDLRKTITIIIPDDPLTF 1380
QY 1381 SSGSRMLNDPSSQSDSEEWMTSLBLAHKDFVSALPDLDBHCACGGENLSVGOBOLVCL 1440
DB 1381 SSGSRMLNDPSSQSDSEEWMTSLBLAHKDFVSALPDLDBHCACGGENLSVGOBOLVCL 1440
QY 1441 ABALLRKTIIIVLDEATAVAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500
DB 1441 ABALLRKTIIIVLDEATAVAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500
QY 1501 DKGEIOEYGAPSDLIQORGLFYSMAXKDGVL 1531
DB 1501 DKGEIOEYGAPSDLIQORGLFYSMAXKDGVL 1531

```

RESULT 8
AA55799
ID AA55799 standard; protein; 1531 AA.

```

AC AAY55799;
XX 28-FEB-2000 (first entry)
DT 28-FEB-2000 (first entry)
XX Human multidrug resistance-associated protein (MRP) variant.
DE Chemosensitizer; multidrug resistance-associated protein; MRP; human;
KW chemosensitizer; multidrug resistance-associated protein; MRP; human;
KW cancer; variant.
XX Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Key 685
FT Misc-difference 685
FT /label= L685S
FT /note= "wild-type Leu is replaced by Ser"
FT Misc-difference 1282
FT /label= R1282A
FT /note= "wild-type Arg is replaced by Ala"
XX US6001563-A.
XX 14-DEC-1999.
XX 05-JUN-1995; 95US-00463179.
XX 27-OCT-1992; 92US-00966923.
XX 08-MAR-1993; 93US-00029340.
XX 26-OCT-1993; 93US-00141893.
XX 20-MAR-1995; 95US-00407207.
XX (TOOH) UNIV QUEBENS KINGSTON.
XX Cole SP, Deeley RG;
XX WPI; 2000-061877/05.
XX DR N-PSDB; AAZ39556.
XX Identification of chemosensitizers useful for treating cancer, using
PT nucleic acid encoding multidrug resistance-associated protein.
XX Claim 3; Col 69-80; 77pp; English.
PS The invention provides a method for identifying a substance which is a
XX chemosensitizer that comprises, contacting a cell transfected with
XX nucleic acid encoding multidrug resistance-associated protein (MRP) with
XX a therapeutic agent in vitro. The method is useful for identifying
XX chemosensitizers which may then be used to treat cancer (especially lung
XX cancer). The method allows the identification of chemosensitizers which
XX do not reverse P-glycoprotein-mediated multidrug resistance. The present
XX sequence represents a human MRP variant
XX Sequence 1531 AA:
SQ
Query Match 79.4%; Score 7860; DB 3; Length 1531;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALMGFCSADSDSLMDMNTVMTNSNDPTKCFQNTVLAWVPCYTLAACPFYFLYLSRH 60
DB 1 MALMGFCSADSDSLMDMNTVMTNSNDPTKCFQNTVLAWVPCYTLAACPFYFLYLSRH 60
QY 61 DRGIYQMTPLNKKTALGFLIMIVCMADLFYSFERSRGIFLAVFLVSPFLIGITITLLA 120
DB 61 DRGIYQMTPLNKKTALGFLIMIVCMADLFYSFERSRGIFLAVFLVSPFLIGITITLLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWALVCAALILRSKIMTALKEDAQVDLFFDITFYVYFS 180
DB 121 TFLIQLERRKGVQSSGIMLTFWALVCAALILRSKIMTALKEDAQVDLFFDITFYVYFS 180
QY 181 LLLIQLVLSGFSDBSPLFSEITIHDPNCPBSSASFLSRTTWTATGLIVRGYRQPLBESD 240
DB 181 LLLIQLVLSGFSDBSPLFSEITIHDPNCPBSSASFLSRTTWTATGLIVRGYRQPLBESD 240

```

QY 241 LMSLNKEDTSEOVVPLVWKWKKECAKTRKOPVKVYSSKDPAPQKESKVDANEVEAL 300
 DB 241 LMSLNKEDTSEOVVPLVWKWKKECAKTRKOPVKVYSSKDPAPQKESKVDANEVEAL 300
 QY 301 IVKSPQKEMNPSLFKVLKFTGFPFLMSFFPKA1HDLMMFSGPOL1K1FVNDTPAD 360
 DB 301 IVKSPQKEMNPSLFKVLKFTGFPFLMSFFPKA1HDLMMFSGPOL1K1FVNDTPAD 360
 QY 361 WQGFYVVLFFVACLOTLVHOFHICFVSGNRIKTAIVGAVRKALVITNSARKSTV 420
 DB 361 WQGFYVVLFFVACLOTLVHOFHICFVSGNRIKTAIVGAVRKALVITNSARKSTV 420
 QY 421 GEIYNLMSVDAORFMDLATTYINM1WSA1PLOYTALTYLMLN1GSPVLAVGAVWMLAMPVN 480
 DB 421 GEIYNLMSVDAORFMDLATTYINM1WSA1PLOYTALTYLMLN1GSPVLAVGAVWMLAMPVN 480
 QY 481 AVNAMKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYAMELAFKDKVLA1ROBELKVLK 540
 DB 481 AVNAMKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYAMELAFKDKVLA1ROBELKVLK 540
 QY 541 KSAVLAVGTFTWCTPPLVALCTPRAVYVITDENN11DAQTA1FVSLAFN1LREP1N1LP 600
 DB 541 KSAVLAVGTFTWCTPPLVALCTPRAVYVITDENN11DAQTA1FVSLAFN1LREP1N1LP 600
 QY 601 MVISSIVQASVSLKRLIFLSHELEBDS1ERRPVKGGGNS1TVNNAFTTARSDPT 660
 DB 601 MVISSIVQASVSLKRLIFLSHELEBDS1ERRPVKGGGNS1TVNNAFTTARSDPT 660
 QY 661 LNCITFS1PBGALVAVVGVCCKSSLSLAL1AMDVEGVAKSGVA1VPOQAW1QND 720
 DB 661 LNCITFS1PBGALVAVVGVCCKSSLSLAL1AMDVEGVAKSGVA1VPOQAW1QND 720
 QY 721 SLREN1LFGCQLEBPYRSV1QACALL1PDL1LPSGDRTE1GEXVNL1SGGQKQVSLAR 780
 DB 721 SLREN1LFGCQLEBPYRSV1QACALL1PDL1LPSGDRTE1GEXVNL1SGGQKQVSLAR 780
 QY 781 AAVSNMDIYLPDPLS1VDAHVGKHTFENY1GPKGMLKXNTR1LYTHSMN1LPQVDY1Y 840
 DB 781 AAVSNMDIYLPDPLS1VDAHVGKHTFENY1GPKGMLKXNTR1LYTHSMN1LPQVDY1Y 840
 QY 841 MSGGKISEMSYQEL1ARDAFAEFLRTYASTBOBOAENG1TVGSGPGKEAKOMENG 900
 DB 841 MSGGKISEMSYQEL1ARDAFAEFLRTYASTBOBOAENG1TVGSGPGKEAKOMENG 900
 QY 901 LVTDSAGKOLQOROLSSSSSYSGDI1SRHNS1TAELQAKAEKKEETWKLMEADKAQTQVKL 960
 DB 901 LVTDSAGKOLQOROLSSSSSYSGDI1SRHNS1TAELQAKAEKKEETWKLMEADKAQTQVKL 960
 QY 961 SYVMDVYKATIGLFTSPLSTFLFMCNHNVSALASNYWLSLMTDDP1VNGTOHETKRLSVYG 1020
 DB 961 SYVMDVYKATIGLFTSPLSTFLFMCNHNVSALASNYWLSLMTDDP1VNGTOHETKRLSVYG 1020
 QY 1021 ALGISOGIAVFGYSMAVISGII1ASRC1AHD1LHS1IRS1PMS1FERT1P1SGNLVNR1FSKEL 1080
 DB 1021 ALGISOGIAVFGYSMAVISGII1ASRC1AHD1LHS1IRS1PMS1FERT1P1SGNLVNR1FSKEL 1080
 QY 1081 DTVDSM1PEVIK1MFMGSLFNV1GACIV1LATP1A1I1PPLG1YFVFORFYVASSROL 1140
 DB 1081 DTVDSM1PEVIK1MFMGSLFNV1GACIV1LATP1A1I1PPLG1YFVFORFYVASSROL 1140
 QY 1141 KRL1ESVRSR1PVYSHFN1TL1GVSV1RAF1EEOER1IHOSDLKVDENQAA1VPS1VANRMA 1200
 DB 1141 KRL1ESVRSR1PVYSHFN1TL1GVSV1RAF1EEOER1IHOSDLKVDENQAA1VPS1VANRMA 1200
 QY 1201 VRL1ECVNC1VL1PAAL1FAV1SRHS1LSAG1VL1SVYS1LQV1TY1LML1VRS1SEMET1VA 1260
 DB 1201 VRL1ECVNC1VL1PAAL1FAV1SRHS1LSAG1VL1SVYS1LQV1TY1LML1VRS1SEMET1VA 1260
 QY 1261 VER1LKEYS1ET1EKA1PMO1Q1ETAP1PSS1PQV1GRV1EFN1Y1CL1RYRED1DL1V1RH1NVT1NGG 1320
 DB 1261 VER1LKEYS1ET1EKA1PMO1Q1ETAP1PSS1PQV1GRV1EFN1Y1CL1RYRED1DL1V1RH1NVT1NGG 1320

QY 1321 EKVGIIVRTGAGKSSLT1LGF1RINESAGE111IDGINIAK1GHLHDLRFKTI1IPQDPVL 1380
 DB 1321 EKVGIIVRTGAGKSSLT1LGF1RINESAGE111IDGINIAK1GHLHDLRFKTI1IPQDPVL 1380
 QY 1381 SSG1LRN1LDPFSGYSDEE1VTS1ELAH1KDFVSA1LDPK1LHCECAEGEN1SVGQROLVCL 1440
 DB 1381 SSG1LRN1LDPFSGYSDEE1VTS1ELAH1KDFVSA1LDPK1LHCECAEGEN1SVGQROLVCL 1440
 QY 1441 ARA1LRKTK1LV1DEATAAD1LETDD1LOST1TQFEDCTV1TAHRLNT1MDTV1VL 1500
 DB 1441 ARA1LRKTK1LV1DEATAAD1LETDD1LOST1TQFEDCTV1TAHRLNT1MDTV1VL 1500
 QY 1501 DKGE1OEYGA1PSDL1QORGL1FYSMAKDA1GLV 1531
 DB 1501 DKGE1OEYGA1PSDL1QORGL1FYSMAKDA1GLV 1531

RESULT 9
 AA78873
 ID AA78873 standard; protein; 1531 AA.
 XX AA78873;
 AC AA78873;
 XX 19-MAY-2000 (first entry)
 DT 19-MAY-2000 (first entry)
 XX
 DE Multidrug resistance protein (MRP) natural variant amino acid sequence.
 XX
 KM Multidrug resistance protein; MRP; human; anthracycline; Vinca alkaloid;
 KW epipodophyllotoxin; cancer; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN US6025473-A.
 XX 15-FEB-2000.
 PD 15-FEB-2000.
 XX
 PF 05-JUN-1995; 95US-00461384.
 PR 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX
 PA (TMOH) UNIV QUEBENS KINGSTON.
 XX
 PI Cole SPC, Deeley RG;
 PT WPI; 2000-181838/16.
 DR N-PDB; AA290193.
 XX
 XX Isolated protein conferring multidrug resistance, to at least two drugs
 PT selected from anthracyclines, epipodophyllotoxins and Vinca alkaloids, on
 PT a drug sensitive mammalian cell.
 XX
 PS Claim 10; Col 79-88; 78pp; English.
 XX
 XX This sequence represents a human multidrug resistance protein (MRP)
 CC natural variant amino acid sequence. The human MRP confers multidrug
 CC resistance, including resistance to at least two drugs selected from
 CC anthracyclines, epipodophyllotoxins and Vinca alkaloids, on a drug
 CC sensitive mammalian cell, when the protein is expressed in the cell. The
 CC multidrug resistance is not substantially reversed by chemosensitizers
 CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP
 CC protein sequence can be used to generate antibodies against MRP. The MRP
 CC protein and nucleotide sequences can be used in compositions which are
 CC used to treat patients with tumours displaying multidrug resistance. The
 CC compositions and methods of the invention can be used particularly to
 CC treat breast cancer, leukaemia, fibrosarcomas, cervical cancer, gliomas,
 CC thymoma, neuroblastomas, and lung cancer. Antibodies directed against
 CC MRP can be used to inhibit the multidrug resistance of a multidrug
 CC resistant cell
 XX
 XX Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 3; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALRGFCSAGSDLDLMDNMTNTNSNDPTFCQNTLVWVPCFYTLMAACFFPYLYLSRH 60
DB 1 MALRGFCSAGSDLDLMDNMTNTNSNDPTFCQNTLVWVPCFYTLMAACFFPYLYLSRH 60
QY 61 DRGIQNTPLINKTALGFLLMIYCNADLFYSFMRSGIFLAVFLVPSPLLGITLLA 120
DB 61 DRGIQNTPLINKTALGFLLMIYCNADLFYSFMRSGIFLAVFLVPSPLLGITLLA 120
QY 121 TFLIQLRRKGVSSGIMLTFWLVALYCALAIILRSKIMTALKEDAQVDLDFDITFYVYS 180
DB 121 TFLIQLRRKGVSSGIMLTFWLVALYCALAIILRSKIMTALKEDAQVDLDFDITFYVYS 180
QY 181 LLLIQLVLSCTSDSPLEFSETHDPNCPRESSASFLSITFWMTGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSCTSDSPLEFSETHDPNCPRESSASFLSITFWMTGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKMKKCECAKTRKQPVKVVYSSKDPAPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKMKKCECAKTRKQPVKVVYSSKDPAPKSSKYDANEVEAL 300
QY 301 IVKSPOKEMNBSLFKVLKTEGPFYFLMSFFPKAIHDLMMFSGPOLKLLIKFVNDTRAPD 360
DB 301 IVKSPOKEMNBSLFKVLKTEGPFYFLMSFFPKAIHDLMMFSGPOLKLLIKFVNDTRAPD 360
QY 361 WQGFYVVLVFNACLOTLVHQYFHICFVSGMRKKTAVIGAVRKALVITNSARKSTV 420
DB 361 WQGFYVVLVFNACLOTLVHQYFHICFVSGMRKKTAVIGAVRKALVITNSARKSTV 420
QY 421 GEIYNLSVDAPQFMDLATYINMIWSAPLOVILALYLLMLNGPSVLGAVAVMLVNVN 480
DB 421 GEIYNLSVDAPQFMDLATYINMIWSAPLOVILALYLLMLNGPSVLGAVAVMLVNVN 480
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILLNGIKVYLKLYAMELAFKDYLAIROBELKYLK 540
DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILLNGIKVYLKLYAMELAFKDYLAIROBELKYLK 540
QY 541 KSATLSAVGTFTWCTPFLVALCTPAVYVITDENNILLDAQAFVSLAFNLRPLNLTLP 600
DB 541 KSATLSAVGTFTWCTPFLVALCTPAVYVITDENNILLDAQAFVSLAFNLRPLNLTLP 600
QY 601 MVISIVOASVSLKRLRFLSHBELEPDSIRRPVKDGGTNSITVRNATFTMARSDPT 660
DB 601 MVISIVOASVSLKRLRFLSHBELEPDSIRRPVKDGGTNSITVRNATFTMARSDPT 660
QY 661 LMGITFSIPGALVAVVQVCGKSSLLSALLAEMDKYEGHVAIKGVAVYVPOQAMIOND 720
DB 661 LMGITFSIPGALVAVVQVCGKSSLLSALLAEMDKYEGHVAIKGVAVYVPOQAMIOND 720
QY 721 SLRENIIFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQKQVSLAR 780
DB 721 SLRENIIFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQKQVSLAR 780
QY 781 AVYENADLYLFDLPSAADAHVGHKIFENVIKPKMKLNKRIILVTHSMYLPQVDVITV 840
DB 781 AVYENADLYLFDLPSAADAHVGHKIFENVIKPKMKLNKRIILVTHSMYLPQVDVITV 840
QY 841 MSGGKISEMSGYOELIARDGAFAEFLRTYASTEQDADAEENGVTGVSFGKEAKOMENG 900
DB 841 MSGGKISEMSGYOELIARDGAFAEFLRTYASTEQDADAEENGVTGVSFGKEAKOMENG 900
QY 901 LVTDASAGKOLQROUSSSSSYSGDISRHNSSTAELOKAKKEETWKLMEADKAGQYKL 960
DB 901 LVTDASAGKOLQROUSSSSSYSGDISRHNSSTAELOKAKKEETWKLMEADKAGQYKL 960
QY 961 SVYDYMKAIGLFTSIFLPMCHVSLASNYLSTLMTDPIVNGOEHHTKVALSYG 1020
DB 961 SVYDYMKAIGLFTSIFLPMCHVSLASNYLSTLMTDPIVNGOEHHTKVALSYG 1020

```

```

QY 1021 ALGISGIANFVGSMAVSGIILASRCLHYDLHSILRSMSPPERTPSGNLVNREKEL 1080
DB 1021 ALGISGIANFVGSMAVSGIILASRCLHYDLHSILRSMSPPERTPSGNLVNREKEL 1080
QY 1081 DTDSMTPEVIKMFMSLENVIGACIVILATPAIIIPPLGIYFVYORFYASSROL 1140
DB 1081 DTDSMTPEVIKMFMSLENVIGACIVILATPAIIIPPLGIYFVYORFYASSROL 1140
QY 1141 KRLESVRSFVYSHFNELLGVSIVIRAFEEQERFIHOSDKVDENQAYPSIVANRMLA 1200
DB 1141 KRLESVRSFVYSHFNELLGVSIVIRAFEEQERFIHOSDKVDENQAYPSIVANRMLA 1200
QY 1201 VRLECVNCIVLPALPAVYSRHSLSAGLVGSYSLSQVTTYLNMLVRMSSEMETIVA 1260
DB 1201 VRLECVNCIVLPALPAVYSRHSLSAGLVGSYSLSQVTTYLNMLVRMSSEMETIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPPSSMPQVGRVFRNYCLRREDDLFVLRHIVTINGG 1320
DB 1261 VERLKEYSETEKEAPMOIOETAPPSSMPQVGRVFRNYCLRREDDLFVLRHIVTINGG 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVL 1380
DB 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVL 1380
QY 1381 SSGIRANLDPSSQYSDDEEVTSLFLAHLKQFVSLPKLHCEKAGEENLSVGQROLVCL 1440
DB 1381 SSGIRANLDPSSQYSDDEEVTSLFLAHLKQFVSLPKLHCEKAGEENLSVGQROLVCL 1440
QY 1441 ABALRKTKILVUDEAFAVADLETDDLIOSTIRPOFEDCVLTATRAHNTIMDTRYIVL 1500
DB 1441 ABALRKTKILVUDEAFAVADLETDDLIOSTIRPOFEDCVLTATRAHNTIMDTRYIVL 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYSMAXDAGLV 1531
DB 1501 DKGEIOEYGAPSDILQORGLFYSMAXDAGLV 1531

```

RESULT 10

ABG61810
 ID ABG61810 standard; protein; 1531 AA.

XX AC ABG61810;
 XX DT 15-AUG-2002 (first entry)
 XX DE

XX Prostate cancer-associated protein #11.

XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX Mammalia.

XX OS
 XX PN WO200230268-A2.

XX PD 18-APR-2002.

XX PF 12-OCT-2001; 2001WO-US032045.

XX PR 13-OCT-2000; 2000US-00687576.

XX PR 08-DEC-2000; 2000US-00733288.

XX PR 08-DEC-2000; 2000US-00733742.

XX PR 24-JAN-2001; 2001US-0263957P.

XX PR 16-MAR-2001; 2001US-0276791P.

XX PR 06-APR-2001; 2001US-0281922P.

XX PR 24-APR-2001; 2001US-0286214P.

XX PR 30-APR-2001; 2001US-00847046.

XX PR 04-MAY-2001; 2001US-0288589P.

XX PA (BOSB-) BOS BIOTECHNOLOGY INC.

XX PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
 XX WP1; 2002-471335/50.

DR N-PSDB; ABK92125.
XX Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
XX Claim 27; Page 309; 436pp; English.
XX
XX The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC gene) that selectively
CC hybridize to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
CC
XX
SQ Sequence 1531 AA;
Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFSAQSDSDPLMDMNTMTNTSNDPTKQNTVLMVPCGYLMACPFYFLYSRH 60
DB 1 MALRGFSAQSDSDPLMDMNTMTNTSNDPTKQNTVLMVPCGYLMACPFYFLYSRH 60
QY 61 DRGVIQMTPLNKTKTALGFLIMIVCMADLFYSFMERSRGIFLAFVFLSPILGITLLA 120
DB 61 DRGVIQMTPLNKTKTALGFLIMIVCMADLFYSFMERSRGIFLAFVFLSPILGITLLA 120
QY 121 TFLIQLERRRKGVOSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLFRDITFYVFS 180
DB 121 TFLIQLERRRKGVOSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLFRDITFYVFS 180
QY 181 LLLIQLVLSCESSDPLSETIHDNCPRESSASFLSRITFWMTTGLIVRGYRPLGSD 240
DB 181 LLLIQLVLSCESSDPLSETIHDNCPRESSASFLSRITFWMTTGLIVRGYRPLGSD 240
QY 241 LMSLNKEDTSEQVVPVVKWKKECAKTRKQPVKVYSSKQPAQPKSSKVDANEYEAL 300
DB 241 LMSLNKEDTSEQVVPVVKWKKECAKTRKQPVKVYSSKQPAQPKSSKVDANEYEAL 300
QY 301 IVKSPQKEMNPSEFKVLYKTFGPFYLMSPFKAIHDLMPGSGPQILKLIFVNDTRAP 360
DB 301 IVKSPQKEMNPSEFKVLYKTFGPFYLMSPFKAIHDLMPGSGPQILKLIFVNDTRAP 360
QY 361 WQGFYTVLLFVTAQOTLVLAHQYFHCFFVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTVLLFVTAQOTLVLAHQYFHCFFVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQREMDLATYINMISAPLOVILALYLLMLNIGSPVLAGVAVMLAMPVN 480
DB 421 GEIYNLMSVDAQREMDLATYINMISAPLOVILALYLLMLNIGSPVLAGVAVMLAMPVN 480
QY 481 AVNAKTKTYQVAHMSKDNRIKIMNEILNGIKVILKYAMELAFKDYALIRQELKVLK 540
DB 481 AVNAKTKTYQVAHMSKDNRIKIMNEILNGIKVILKYAMELAFKDYALIRQELKVLK 540
QY 541 KSAVLASVGTFTWCTPEIVALCTFAVYVITDENNNIIDAGTAFLSLFNLIRPLNLP 600
DB 541 KSAVLASVGTFTWCTPEIVALCTFAVYVITDENNNIIDAGTAFLSLFNLIRPLNLP 600
QY 601 MVISSIYQASVSLKRLIFLSHEELPDSIERRPVKQGGGNTSYVNAFTMARSDPT 660
DB 601 MVISSIYQASVSLKRLIFLSHEELPDSIERRPVKQGGGNTSYVNAFTMARSDPT 660

DB 601 MVISSIYQASVSLKRLIFLSHEELPDSIERRPVKQGGGNTSYVNAFTMARSDPT 660
QY 661 LNGITTSIPGALVAVYGVGCCGSSLLSALLAMDMYEGHVAIKGSVAVYPOQAWIQND 720
DB 661 LNGITTSIPGALVAVYGVGCCGSSLLSALLAMDMYEGHVAIKGSVAVYPOQAWIQND 720
QY 721 SLRENIIFGCOLSEPPYRSVYIQAACALPDLEILPDSGRTEIGEGVNLSSGQKQVSLAR 780
DB 721 SLRENIIFGCOLSEPPYRSVYIQAACALPDLEILPDSGRTEIGEGVNLSSGQKQVSLAR 780
QY 781 AVYSNADIVYFDDELSAVDAHVGKHIPENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYSNADIVYFDDELSAVDAHVGKHIPENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGGKIEMSGSYQELARDDGAPAFLLTYASTEBQDAEENGVTGSGPGKEAQMENG 900
DB 841 MSGGKIEMSGSYQELARDDGAPAFLLTYASTEBQDAEENGVTGSGPGKEAQMENG 900
QY 901 LVTDASGKOLOROLSSSSYSGDISRHNSTAELOKAEKKEETWKLMEADKATQGVKL 960
DB 901 LVTDASGKOLOROLSSSSYSGDISRHNSTAELOKAEKKEETWKLMEADKATQGVKL 960
QY 961 SVYWDYMKALGLFISFLSIFLFCMCHVSALASNYMTSLMTDDPIVNGTOEHTKVLSTYVG 1020
DB 961 SVYWDYMKALGLFISFLSIFLFCMCHVSALASNYMTSLMTDDPIVNGTOEHTKVLSTYVG 1020
QY 1021 ALGISQIAVFGYSMAVSIQIILASRCLAVDLHSIIRSPMSFEPTPSGNLVNRSKEL 1080
DB 1021 ALGISQIAVFGYSMAVSIQIILASRCLAVDLHSIIRSPMSFEPTPSGNLVNRSKEL 1080
QY 1081 DTWDSMIBEVIKPMFMSLFENVIGACIVILATPAIILPILGITYFNQRFVYASSROL 1140
DB 1081 DTWDSMIBEVIKPMFMSLFENVIGACIVILATPAIILPILGITYFNQRFVYASSROL 1140
QY 1141 KRLSEYSRSPVYSHFNFTLLGVSIVIRAFEEQERFIHOSDLKVDENQKAYPSIVANWLA 1200
DB 1141 KRLSEYSRSPVYSHFNFTLLGVSIVIRAFEEQERFIHOSDLKVDENQKAYPSIVANWLA 1200
QY 1201 VRLECVNGCIYLPALFAVISRHSLSAGLVGSVSYSLQVTTYLNLVRRSMSEMETNIVA 1260
DB 1201 VRLECVNGCIYLPALFAVISRHSLSAGLVGSVSYSLQVTTYLNLVRRSMSEMETNIVA 1260
QY 1261 VERLKEYSERTEKAPMIOETAPSSWPQVGRFEFRNYCLRYREDLDPVLRHNVTTING 1320
DB 1261 VERLKEYSERTEKAPMIOETAPSSWPQVGRFEFRNYCLRYREDLDPVLRHNVTTING 1320
QY 1321 EKYGIQRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKYGIQRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
QY 1381 SSGILRNMLDPFSQYSDEEVTSLLEHLKDFVSALPDKLDHECABEGENISVGOQLVCL 1440
DB 1381 SSGILRNMLDPFSQYSDEEVTSLLEHLKDFVSALPDKLDHECABEGENISVGOQLVCL 1440
QY 1441 ARALLRRTKTLVDEATAAVALDETDDLIOGTITRQEDCTVTLIARHLANTIMYTRIVL 1500
DB 1441 ARALLRRTKTLVDEATAAVALDETDDLIOGTITRQEDCTVTLIARHLANTIMYTRIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
RESULT 11
ABM35012
ID ABM35012 standard; protein; 1531 AA.
AC ABM35012;
XX
XX 08-OCT-2003 (first entry)
XX
XX Cancer based on CYP3A5 related protein SEQ ID NO:678.
DE

KW Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;
KM cytochrome p450; subfamily IIA; nifedipine oxidase; polypeptide 5;
XX cytochrome p450.
XX
XX Unidentified.
XX
XX WO2003013534-A2.
XX
XX 20-FEB-2003.
XX
XX 23-JUL-2002; 2002WO-EP008219.
XX
XX 23-JUL-2001; 2001EP-00117608.
XX
XX 24-MAY-2002; 2002EP-00011710.
XX
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
XX Heinrich G, Kerb R,
XX
XX WPI; 2003-268144/26.
XX
XX
XX New use of irinotecan for preparation of compositions for treating cancer
PT in subject having genome with variant allele comprising cytochrome p450,
PT subfamily IIA, polypeptide 5 polynucleotide, termed CYP3A5.
XX
XX Disclosure; SEQ ID NO 678; 86pp; English.
XX
XX
XX The present invention describes the use of irinotecan (I) or its
XX derivative for the preparation of a pharmaceutical composition for
XX treating colorectal, cervical, gastric, lung, ovarian or pancreatic
XX cancer, or malignant glioma in a subject having a genome with a variant
XX allele which comprises a cytochrome p450, subfamily IIA (nifedipine
XX oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have
XX cytochrome activity. The therapeutic applications of (I) is improved,
XX since it is possible to individually treat a subject with an appropriate
XX dosage and/or an appropriate derivative of (I). Therefore, undesirable,
XX harmful or toxic effects are efficiently avoided. Unnecessary and
XX potentially harmful treatment of those subjects who do not respond to the
XX treatment with substances (nonresponders), as well as the development of
XX drug resistances due to suboptimal drug dosing can be avoided. ACF62200
XX to ACF622751 and ABM34912 to ABM35013 represent sequences used in the
XX exemplification of the present invention
XX
XX
XX Sequence 1531 AA;
SQ
Query Match 79.4%; Score 7860; DB 6; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALNGFCSDGSDPLMDNMTNTNSPDTKCFONTVLWVPCFYLMACPPEFYFLYSRH 60
DB 1 MALNGFCSDGSDPLMDNMTNTNSPDTKCFONTVLWVPCFYLMACPPEFYFLYSRH 60
QY 61 DRGIQMTPLNKTATLGPLLMTVCWADLFYSFWRSGIFLAPVFLVSPITLGLITTLA 120
DB 61 DRGIQMTPLNKTATLGPLLMTVCWADLFYSFWRSGIFLAPVFLVSPITLGLITTLA 120
QY 121 TFLQLERRKGVQSSGIMLTFPLVALVCAALILRSKIMTALKEQVNDLFRITTYVVS 180
DB 121 TFLQLERRKGVQSSGIMLTFPLVALVCAALILRSKIMTALKEQVNDLFRITTYVVS 180
QY 181 LLLIQLVLSCFSDRSLPSETIHDNPPCESASFLSRITTFWMIIGLIVRGYRQPLESSD 240
DB 181 LLLIQLVLSCFSDRSLPSETIHDNPPCESASFLSRITTFWMIIGLIVRGYRQPLESSD 240
QY 241 LMSLNKEDTSSQVVPVLVKNMKKECAKTRKQPVKVVYSKODAPQKSSKVDANEVVAL 300
DB 241 LMSLNKEDTSSQVVPVLVKNMKKECAKTRKQPVKVVYSKODAPQKSSKVDANEVVAL 300
QY 301 IVKSQKEMNBSLFLVLYKTGRPYFLMFPFKAIIIDLMFSGPOLIKLIRVNTTKAPD 360
DB 301 IVKSQKEMNBSLFLVLYKTGRPYFLMFPFKAIIIDLMFSGPOLIKLIRVNTTKAPD 360

QY 361 WQGYFTVLLFVTAQTLVLAHQYFHCYVSGMKRKTAVGAVYRKALVITNSARKSSTV 420
DB 361 WQGYFTVLLFVTAQTLVLAHQYFHCYVSGMKRKTAVGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRMDLATYINMTWSAPLOVITLALYLMNIGPSVLAGVAVWVLAHPVN 480
DB 421 GEIVNLSVDAQRMDLATYINMTWSAPLOVITLALYLMNIGPSVLAGVAVWVLAHPVN 480
QY 481 AVNMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVILKLYAMELAFKDYVALIROLELKYLK 540
DB 481 AVNMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVILKLYAMELAFKDYVALIROLELKYLK 540
QY 541 KSAVLSAVGFTFWCTPFLVALCTPANYVTIIDENNIIIDAOTAFYSLAFNLRPLNLP 600
DB 541 KSAVLSAVGFTFWCTPFLVALCTPANYVTIIDENNIIIDAOTAFYSLAFNLRPLNLP 600
QY 601 MVTSIYQASVSLKRLNIFLSHELEPDSIERRPVKGGGNSITVNAFTTARSPT 660
DB 601 MVTSIYQASVSLKRLNIFLSHELEPDSIERRPVKGGGNSITVNAFTTARSPT 660
QY 661 LMGITFSPREGALVAVVGVCGCKSSLALSALLAEMDRVGEHVAIKGSVAVYVQQAWIQND 720
DB 661 LMGITFSPREGALVAVVGVCGCKSSLALSALLAEMDRVGEHVAIKGSVAVYVQQAWIQND 720
QY 721 SLRENIIIFGCOLLEPYRVSIVQACALLPDLEILPSGDRTEIGEGVNLSGQKRVSLAR 780
DB 721 SLRENIIIFGCOLLEPYRVSIVQACALLPDLEILPSGDRTEIGEGVNLSGQKRVSLAR 780
QY 781 AVYENADIVLEPDBLSAVDAHVGHIFENYIGPMKNKTRILVTHSMSTLPQVDYIV 840
DB 781 AVYENADIVLEPDBLSAVDAHVGHIFENYIGPMKNKTRILVTHSMSTLPQVDYIV 840
QY 841 MSGGKISEMGSYQELLARDGAFALFRTYASTBOEODAEENGTVGSGPGKEAKOMENG 900
DB 841 MSGGKISEMGSYQELLARDGAFALFRTYASTBOEODAEENGTVGSGPGKEAKOMENG 900
QY 901 LVYDSAGKQLOKQUSSSSSYSGDISRHNSHTAEIQAKEAKKEFTWKMEADKAOTGOVKL 960
DB 901 LVYDSAGKQLOKQUSSSSSYSGDISRHNSHTAEIQAKEAKKEFTWKMEADKAOTGOVKL 960
QY 961 SVYDMYKATGLFTSFISTFLFMCNHSALASNVWLSLMTDDPLVNGQHTKRLSYG 1020
DB 961 SVYDMYKATGLFTSFISTFLFMCNHSALASNVWLSLMTDDPLVNGQHTKRLSYG 1020
QY 1021 ALGISOGIAVGYMAVSIIGIILASRCHVDLHLSILSPMSFEERTPSGNLVNRSKEL 1080
DB 1021 ALGISOGIAVGYMAVSIIGIILASRCHVDLHLSILSPMSFEERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPEVIKMGMSLFNVIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
DB 1081 DTVDSMIPEVIKMGMSLFNVIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
QY 1141 KRLSSVRSPPYSHFNFTLGVSYIYAFEEBERITHOSDLKVDENOKAYYSIVANRPLA 1200
DB 1141 KRLSSVRSPPYSHFNFTLGVSYIYAFEEBERITHOSDLKVDENOKAYYSIVANRPLA 1200
QY 1201 VRLCEVNGCIYLFALFAVISRHSLSAGVGLSVYSLOVTTYNMVLRMSSEMETNIVA 1260
DB 1201 VRLCEVNGCIYLFALFAVISRHSLSAGVGLSVYSLOVTTYNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEYSETBEKAPWQIOETAPPSSWPQVRVEFRNYCLAREYRDLDFLHINVTINGG 1320
DB 1261 VERLKEYSETBEKAPWQIOETAPPSSWPQVRVEFRNYCLAREYRDLDFLHINVTINGG 1320
QY 1321 EKVGIYGTGAGKSSLTGLFRINESAGEIINGINIAKIGLHDLRKTITTIQDPPLF 1380
DB 1321 EKVGIYGTGAGKSSLTGLFRINESAGEIINGINIAKIGLHDLRKTITTIQDPPLF 1380
QY 1381 SGLRMLNDPFSQXSDEEWTSLTAHLKQDVSLAPDLDBEACBAGGENLSVGOROLVCL 1440
DB 1381 SGLRMLNDPFSQXSDEEWTSLTAHLKQDVSLAPDLDBEACBAGGENLSVGOROLVCL 1440
QY 1441 ARALLRKTILVLEDEATAVDLETTDDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500

Db 1441 ARALLRKTILVLEDEATAVDELTDLQSTIRTRQFEQCTVLTIAHRLNTIMDYRIVL 1500

Qy 1501 DKGEIQEYGASDILQORGLFYMAKXAGLV 1531
Db 1501 DKGEIQEYGASDILQORGLFYMAKXAGLV 1531

RESULT 12

ID ADB20865 standard; protein; 1531 AA.

AC ADB20865;

DT 20-NOV-2003 (first entry)

DE MRP1 based cancer related protein SEQ ID NO:678.

KM irinotecan; colorectal cancer; cervical cancer; gastric cancer;

KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;

XX variant allele; multidrug resistance protein 1; MRP1; cytosolic.

OS Unidentified.

PN W02003013533-A2.

PD 20-FEB-2003.

PF 23-JUL-2002; 2002WO-EP008200.

PR 23-JUL-2001; 2001EP-00117608.

PR 24-MAY-2002; 2002EP-00011710.

PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

PI Heinrich G, Kerb R;

XX WPI; 2003-354397/33.

PT Use of irinotecan or its derivative for preparation of a pharmaceutical

PT variant allele comprising a multidrug resistance protein 1

XX polynucleotide.

PS Disclosure; SEQ ID NO 678; 100pp; English.

XX The present invention describes a method for the use of irinotecan (I) or

CC its derivative for the preparation of a pharmaceutical composition for

CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic

CC cancer, or malignant glioma in a subject having a genome with a variant

CC allele which comprises a multidrug resistance protein 1 (MRP1)

CC polynucleotide (II). (I) has cytostatic activity. (II) or its derivative

CC can be used for the preparation of a pharmaceutical composition for

CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic

CC cancer, or malignant glioma in a subject, where the subject is a human

CC (preferably African or Asian) or a mouse. The present sequence represents

CC a sequence which is used in the exemplification of the present invention.

XX Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 6; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDMNTWNTSNPDKFQNTVLYWVPCFYLMACPFYFLYLSRH 60

Db 1 MALRGFCSADGSDPLMDMNTWNTSNPDKFQNTVLYWVPCFYLMACPFYFLYLSRH 60

Qy 61 DRGVIQMTPLNKTITAGFLIMTYCWDLFYSFWERSGITFLAVFLVSPILGITTLLA 120

Db 61 DRGVIQMTPLNKTITAGFLIMTYCWDLFYSFWERSGITFLAVFLVSPILGITTLLA 120

Db 121 TFLQERRRGVQSSGIMLTFWLVALVCALAILSKINTALKEDAOVDLFRDITFYVYS 180

Qy 181 LLLIQLVISCESDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLIVRGYRPLEGSD 240

Db 181 LLLIQLVISCESDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLIVRGYRPLEGSD 240

Qy 241 LMSLNKEDTSQVVPVLYKMKKECAKTRQPVVYVSSKDPAPKSSSKVDANEVVAL 300

Db 241 LMSLNKEDTSQVVPVLYKMKKECAKTRQPVVYVSSKDPAPKSSSKVDANEVVAL 300

Qy 301 IVKSPKEMNPSELFVYKTFEPYFLMSFFPKA IHDLMFSGPOLIKLLIKFVNDTKAPD 360

Db 301 IVKSPKEMNPSELFVYKTFEPYFLMSFFPKA IHDLMFSGPOLIKLLIKFVNDTKAPD 360

Qy 361 WQGYFTVLLFVTACIOTLVHQQYFHI CPVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420

Db 361 WQGYFTVLLFVTACIOTLVHQQYFHI CPVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420

Qy 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOVITALLYLMLNGPSVLAGAVWVLMVPMVN 480

Db 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOVITALLYLMLNGPSVLAGAVWVLMVPMVN 480

Qy 481 AYVAMKTKTYQVAAHMKSKDNRIKLMNEILNGIKVLYAMELAFKDKVLAIROBELKVLK 540

Db 481 AYVAMKTKTYQVAAHMKSKDNRIKLMNEILNGIKVLYAMELAFKDKVLAIROBELKVLK 540

Qy 541 KSAVLSAVGTFTWVCTPFVVALCTFAVYVITDENNIIDAOFAFVSLFNLIRPPLNLP 600

Db 541 KSAVLSAVGTFTWVCTPFVVALCTFAVYVITDENNIIDAOFAFVSLFNLIRPPLNLP 600

Qy 601 MVISSIVQASVSLKRLIFLSHELEBDSIERRVKQGGGNSITVFNATFTWARSPT 660

Db 601 MVISSIVQASVSLKRLIFLSHELEBDSIERRVKQGGGNSITVFNATFTWARSPT 660

Db 1201 VRLECVNGNCIVLPALPAVAVIRSHSLAGLVGLSVSYSLQVTTYLNLVVRMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320
Qy 1321 EKVGIVGTGTAGKSSLTGLFRINBSAGEEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
Db 1321 EKVGIVGTGTAGKSSLTGLFRINBSAGEEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
Qy 1381 SGLSRMNLDPSSQVSDSEVMTSLSLAHKDFVSALPDLDHFCAGGSENLVSGORQIVCL 1440
Db 1381 SGLSRMNLDPSSQVSDSEVMTSLSLAHKDFVSALPDLDHFCAGGSENLVSGORQIVCL 1440
Qy 1441 ARALLRKTKIIVLDEATAVLDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKTKIIVLDEATAVLDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Qy 1501 DKGEIOERYGAPSDLIQORGLFTYSMAKXAGLV 1531
Db 1501 DKGEIOERYGAPSDLIQORGLFTYSMAKXAGLV 1531

RESULT 13

ADB87954 ID ADB87954 standard; protein; 1531 AA.

XX ADB87954;

DT 04-DEC-2003 (first entry)

XX Human UGT1A1 protein sequence SEQ ID NO:678.

XX irinotecan; cancer; UGT1A1; cytosolic; topoisomerase I inhibitor;
KM colorectal cancer; cervical cancer; gastric cancer; lung cancer;
KM ovarian cancer; pancreatic cancer; malignant glioma;
KM uridine diphosphate glycosyltransferase1 member A1.

XX Homo sapiens.

XX WO2003013536-A2.

XX PD 20-FEB-2003.

XX PF 23-JUL-2002; 2002WO-EP008217.

XX PR 23-JUL-2001; 2001EP-00117608.

XX PR 24-MAY-2002; 2002EP-00011710.

XX PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX PI Heinrich G, Kerb R;

XX WPI; 2003-289896/28.

PT Use of irinotecan to treat cancer patient by determining if patient has
PT variant alleles of UGT1A1 gene, administering increased/decreased amounts
PT of irinotecan based on increased/decreased levels of UGT1A1 gene product.
XX
XX Disclosure; SEQ ID NO 678; 107bp; English.

XX The invention relates to the novel use of irinotecan to treat a patient
XX suffering from cancer. This involves determining if the patient has one
XX or more variant alleles of the UGT1A1 gene, and if the patient has one
XX or more of such variant alleles, irinotecan is administered in an increased
XX or decreased amount in comparison to the amount that is administered
XX without regard to the patient's alleles in the UGT1A1 gene. The invention
XX has cytosolic activity. A composition of the invention acts as a
XX topoisomerase I inhibitor. The method is useful for treating a patient,
XX an animal e.g. mouse or a human, preferably African or Asian, suffering
XX from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
XX pancreatic cancer or malignant glioma. The present sequence is udes in
XX the exemplification of the invention.

XX SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDMNTWNTSNDFPKCFONTLVVWPCFYLLMACPEFFLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTWNTSNDFPKCFONTLVVWPCFYLLMACPEFFLYLSRH 60
Qy 61 DRGYIOMTPLNKTKTALGFLLMVICMADLFYSFMRBGRGIFLAVFLVSPFLIGITLLA 120
Db 61 DRGYIOMTPLNKTKTALGFLLMVICMADLFYSFMRBGRGIFLAVFLVSPFLIGITLLA 120
Qy 121 TFLIQLERRKGVSSGIMLTFMWALVCAALAIKSKMTALKEDAOVDLFFDITFYVYS 180
Db 121 TFLIQLERRKGVSSGIMLTFMWALVCAALAIKSKMTALKEDAOVDLFFDITFYVYS 180
Qy 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCESSASFLSRITFWITGLIVRGYROPLEGSD 240
Db 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCESSASFLSRITFWITGLIVRGYROPLEGSD 240
Qy 241 LMSLNKEDTSROVPUVKMKKECAKTRKQPVKVYVSSKDPAQKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSROVPUVKMKKECAKTRKQPVKVYVSSKDPAQKSSKYDANEVEAL 300
Qy 301 IVKSPKEMNPSLFKVLKYTFGPFLMSFFPKA IHDLMFSGPOLIKLLIKFVNDTKAPD 360
Db 301 IVKSPKEMNPSLFKVLKYTFGPFLMSFFPKA IHDLMFSGPOLIKLLIKFVNDTKAPD 360
Qy 361 WQGYFYTVLLEVTACLOTVLHQYFHI CFSGMBIKTAIVAGVYRKALVITNSARKSTV 420
Db 361 WQGYFYTVLLEVTACLOTVLHQYFHI CFSGMBIKTAIVAGVYRKALVITNSARKSTV 420
Qy 421 GEIVNLSVDAQRMDLATYINMTWSAPLOYIALYILMLNGSVYLAGVAVMTLMPVN 480
Db 421 GEIVNLSVDAQRMDLATYINMTWSAPLOYIALYILMLNGSVYLAGVAVMTLMPVN 480
Qy 481 AVMAKTKTQVAMKSKDNRIKLMNELINIKYLKLYAMELAFKDKVLA ROEBELKYLK 540
Db 481 AVMAKTKTQVAMKSKDNRIKLMNELINIKYLKLYAMELAFKDKVLA ROEBELKYLK 540
Qy 541 KSAVLSAVGFTWVCTPFLVALCTFAVYVITDENNIIDAQTAVSLAFNLRPEPLNLP 600
Db 541 KSAVLSAVGFTWVCTPFLVALCTFAVYVITDENNIIDAQTAVSLAFNLRPEPLNLP 600
Qy 601 MVISIYQASVSLKRLRIFLSHEBELPDSIRRPVKGCGGINSITVNRATFTWARSDEPT 660
Db 601 MVISIYQASVSLKRLRIFLSHEBELPDSIRRPVKGCGGINSITVNRATFTWARSDEPT 660
Qy 661 LMGITFSPREGALVAVVGVGCGKSSLLSALLAEMDKVGHVALKGSAAVYPOQAMIOND 720
Db 661 LMGITFSPREGALVAVVGVGCGKSSLLSALLAEMDKVGHVALKGSAAVYPOQAMIOND 720
Qy 721 SLRENILFGCOLLEPPYRSVYQACALLPDLLEILPSGRTEIGEGVNLSSGQKQVSLAR 780
Db 721 SLRENILFGCOLLEPPYRSVYQACALLPDLLEILPSGRTEIGEGVNLSSGQKQVSLAR 780
Qy 781 AVYSNADLYLPDDPLSAVDHVGKIFENYVGPCKMLKNKRIILVTHSMSTYLPQDVYIV 840
Db 781 AVYSNADLYLPDDPLSAVDHVGKIFENYVGPCKMLKNKRIILVTHSMSTYLPQDVYIV 840
Qy 841 MSGGISMSGYOELLARDGAFAPFLRYASTBEOBDAEENGUVGVSRGKEAKOMEMGM 900
Db 841 MSGGISMSGYOELLARDGAFAPFLRYASTBEOBDAEENGUVGVSRGKEAKOMEMGM 900
Qy 901 LVYTSAGKQLORODSSSSSYSGDISRHNSYAELOKAEAKKEFTWKLMEADKAOQGVYL 960
Db 901 LVYTSAGKQLORODSSSSSYSGDISRHNSYAELOKAEAKKEFTWKLMEADKAOQGVYL 960
Qy 961 SVYDYMKAIGLFIISLIFLMGNCNHSALASNTWLSMTDPIVNGTOEHTKVALSYG 1020
Db 961 SVYDYMKAIGLFIISLIFLMGNCNHSALASNTWLSMTDPIVNGTOEHTKVALSYG 1020

Db 961 SVYWDYKKAIGLFISIFLFCMCHVSAALSNWYLSMTDDPIVNGTOEHKTVKLSVYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLIRSPSPFERPESGULVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLIRSPSPFERPESGULVNRFSKEL 1080
QY 1081 DTVDMSIPEVIAKMGMSLFENVIGACIVILLATPIAAIIIPPLGIYFVQRFYVASSROL 1140
Db 1081 DTVDMSIPEVIAKMGMSLFENVIGACIVILLATPIAAIIIPPLGIYFVQRFYVASSROL 1140
QY 1141 KRLESVSRSPVYSHFNELLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
Db 1141 KRLESVSRSPVYSHFNELLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
QY 1201 VRLFCVGNCTVLFALPAVIRSRHSLASGLVLSVYSIQVTTYLNWLVKMSSEMETIVA 1260
Db 1201 VRLFCVGNCTVLFALPAVIRSRHSLASGLVLSVYSIQVTTYLNWLVKMSSEMETIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVFRNYCLAREDDLPFLSHINVTNGG 1320
Db 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVFRNYCLAREDDLPFLSHINVTNGG 1320
QY 1321 EKVGIVGRTGAGKSLTGLFRINESABGEIIDIINIAKIGLHDLFPKTIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSLTGLFRINESABGEIIDIINIAKIGLHDLFPKTIIPQDPVLF 1380
QY 1381 SSGSRNMLDPPSQSDSEEWTSLELAHKDFVSALPKLDHECAGGENLSVGQROVCL 1440
Db 1381 SSGSRNMLDPPSQSDSEEWTSLELAHKDFVSALPKLDHECAGGENLSVGQROVCL 1440
QY 1441 ARALLRRTKTLVDEATAVADLETDDLIQSTIRTOFEDCTVLTAAHRLNTIMDTRYIVL 1500
Db 1441 ARALLRRTKTLVDEATAVADLETDDLIQSTIRTOFEDCTVLTAAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEGYGA PSDLLQQRGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEGYGA PSDLLQQRGLFYSMAKDAGLV 1531

RESULT 14
ADB96937 standard; protein; 1531 AA.
XX ADB96937;
AC ADB96937;
XX
DT 04-DEC-2003 (first entry)
XX
De Human MDR1 related protein sequence SEQ ID NO:678.
XX
Kw irinotecan; colorectal cancer; cervical cancer; gastric cancer;
Kw lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
Kw multidrug resistance 1; MDR1; cytosolic; human; Cyp3A5; MRP1; MDR1;
Kw TGP1.
XX
OS Homo sapiens.
XX
PN W02003013537-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002W0-EP008218.
XX
PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-0001710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-268145/26.
XX
PT New use of irinotecan for preparation of pharmaceutical compositions for
treating cancer in subject having genome with variant allele comprising

PT multidrug resistance 1 polynucleotide.
XX
PS Disclosure; SEQ ID NO 678; 130bp; English.
XX
CC The invention relates to the novel use of irinotecan or its derivative
CC for the preparation of pharmaceutical compositions for treating
CC colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or
CC malignant glioma in a subject having a genome with a variant allele which
CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition
CC of the invention has cytotoxic activity. The invention is useful for the
CC preparation of pharmaceutical compositions for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject (preferably human, more preferably African or Asian)
CC or a mouse. The present sequence is used in the exemplification of the
CC invention.
XX
SQ Sequence 1531 AA;
XX
Query Match 79.4%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADSDPLMDMNTWNTSNBDFTKCFQNTVLYWVPCFYLMACEPFYLYLSRH 60
Db 1 MALRGFCSADSDPLMDMNTWNTSNBDFTKCFQNTVLYWVPCFYLMACEPFYLYLSRH 60
QY 61 DRGYQMTPLNKTALGFLMTVCADLFFYFMRSRGIFLAPVFLVSPTLGIITLLA 120
Db 61 DRGYQMTPLNKTALGFLMTVCADLFFYFMRSRGIFLAPVFLVSPTLGIITLLA 120
QY 121 TPLIOLERRKGVSSGIMLFWLVALCALAIIIRSKIMTFLKEDAOVDLPDIFYYVFS 180
Db 121 TPLIOLERRKGVSSGIMLFWLVALCALAIIIRSKIMTFLKEDAOVDLPDIFYYVFS 180
QY 181 LLLIOLVLSGFSRSPFLFSETIHDPNCPRESSASFLSRTFFWITGLIVRGYRPLEGSD 240
Db 181 LLLIOLVLSGFSRSPFLFSETIHDPNCPRESSASFLSRTFFWITGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVKKWKKBCKATRKQPVYVYSSODPQPKSSRYDANEVEAL 300
Db 241 LMSLNKEDTSEOVVPLVKKWKKBCKATRKQPVYVYSSODPQPKSSRYDANEVEAL 300
QY 301 IVKSPQKKNPSSLFKVLKTFPGPYFLMSFFFKAIHDLMPSSGQILKLLKFNVDTRAPD 360
Db 301 IVKSPQKKNPSSLFKVLKTFPGPYFLMSFFFKAIHDLMPSSGQILKLLKFNVDTRAPD 360
QY 361 WOGYFYTVLLFVTAQLOTLVLAHQYFHLCFVSGNRKITAIVGAVYRKALVITNSARKSSTV 420
Db 361 WOGYFYTVLLFVTAQLOTLVLAHQYFHLCFVSGNRKITAIVGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNIMSVDAQRPMDIATYINMISAPLOVIALYILNLGPSVLGAVAVWLVWPN 480
Db 421 GEIVNIMSVDAQRPMDIATYINMISAPLOVIALYILNLGPSVLGAVAVWLVWPN 480
QY 481 ANAMKTKQYOVAMMSKONRITLMEIINGIVLKYAMELAFKQVLAIRQBEKLVK 540
Db 481 ANAMKTKQYOVAMMSKONRITLMEIINGIVLKYAMELAFKQVLAIRQBEKLVK 540
QY 541 KSAVLSAVGFTVWCTPFLVALCTFAVYVYIIDENNLDAOTAFVSLAFNLIPLPLILP 600
Db 541 KSAVLSAVGFTVWCTPFLVALCTFAVYVYIIDENNLDAOTAFVSLAFNLIPLPLILP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERPVDGSGTNSIIVRNATFTWARSDDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERPVDGSGTNSIIVRNATFTWARSDDPT 660
QY 661 LNGITFSIPBGALVAVVGVGCGKSSLSLALAMKVBEGHVAIKSVAVVPOQAMQND 720
Db 661 LNGITFSIPBGALVAVVGVGCGKSSLSLALAMKVBEGHVAIKSVAVVPOQAMQND 720
QY 721 SLRENTLFGCOLBEPYRVSITQACALLPDLIELPSGDRTEIGKGVNLSGGQKORVSLAR 780
Db 721 SLRENTLFGCOLBEPYRVSITQACALLPDLIELPSGDRTEIGKGVNLSGGQKORVSLAR 780

QY 781 AVYSNADIIYLPDDPLSAVDAAVHGKII FENVIGPKGMLNKKTRILLVTHSMSTYLPQVDVILV 840
DB 781 AVYSNADIIYLPDDPLSAVDAAVHGKII FENVIGPKGMLNKKTRILLVTHSMSTYLPQVDVILV 840
QY 841 MSGGKISMSGSYOELIADGAPAEFLRTYASTEQDABENGVTGVSPGKAEKMGEM 900
DB 841 MSGGKISMSGSYOELIADGAPAEFLRTYASTEQDABENGVTGVSPGKAEKMGEM 900
QY 901 LVTSAGKQIQROLSSSSSSYSGDISRHNNSTAELOKAKKBEETWKLMEADKAOQGVKL 960
DB 901 LVTSAGKQIQROLSSSSSSYSGDISRHNNSTAELOKAKKBEETWKLMEADKAOQGVKL 960
QY 961 SVYWDYKAIQIFLISFLSIFLMCNHVSALASNTWLSMTDDPIYNGTOEHTKVALSYYG 1020
DB 961 SVYWDYKAIQIFLISFLSIFLMCNHVSALASNTWLSMTDDPIYNGTOEHTKVALSYYG 1020
QY 1021 ALGISQGIATVFGYMANVSTGGILASRCLHVDLLHSILSPMSFFERTPSGULVNRFSKEL 1080
DB 1021 ALGISQGIATVFGYMANVSTGGILASRCLHVDLLHSILSPMSFFERTPSGULVNRFSKEL 1080
QY 1081 DTVMSPMPEVTKMFMGSLFENVIGACTIVLLATPIAIIIPPLGLIYFFVQRFYVASSRQL 1140
DB 1081 DTVMSPMPEVTKMFMGSLFENVIGACTIVLLATPIAIIIPPLGLIYFFVQRFYVASSRQL 1140
QY 1141 KRLESVSRSPYVSHFNETILGVSVIRAFEBQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYVSHFNETILGVSVIRAFEBQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCTILFALPAFVIRSHSLASGLVGLSVSYSLQVTTYLNMVLRMSEMETNIVA 1260
DB 1201 VRLCEVGNCTILFALPAFVIRSHSLASGLVGLSVSYSLQVTTYLNMVLRMSEMETNIVA 1260
QY 1261 VERLKEVSETEKEAPWQIOETAPPSMPQVGRVEFRNCLKREDDVILRHINVTINGG 1320
DB 1261 VERLKEVSETEKEAPWQIOETAPPSMPQVGRVEFRNCLKREDDVILRHINVTINGG 1320
QY 1321 EKVGIVGTGAGKSLTLGLFRINSAEGEIIIDGINIAKIGLHDLRKFTIIPQDPVLF 1380
DB 1321 EKVGIVGTGAGKSLTLGLFRINSAEGEIIIDGINIAKIGLHDLRKFTIIPQDPVLF 1380
QY 1381 SSGIRPMNIDPPSOVSDEEVTSLBLAKDPYSALPDKLBHCAGGENTLVSQROVLVCL 1440
DB 1381 SSGIRPMNIDPPSOVSDEEVTSLBLAKDPYSALPDKLBHCAGGENTLVSQROVLVCL 1440
QY 1441 ARALLRKTKILVLDATAAVDLETFDDLIOSTIRTOFEBCVTYLLIARLNTIMDYRVIVL 1500
DB 1441 ARALLRKTKILVLDATAAVDLETFDDLIOSTIRTOFEBCVTYLLIARLNTIMDYRVIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMKADAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMKADAGLV 1531
RESULT 15
ADB92128
ID ADB92128 standard; protein; 1531 AA.
XX
AC ADB92128;
DT 04-DEC-2003 (first entry)
XX Human MDRI related protein sequence SEQ ID NO:678.
XX
KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW multidrug resistance 1; MDRI; cytosolic; human; UGT1A1; MRP1; TOP1.
XX
OS Homo sapiens.
XX
PN MO2003013535-A2.
XX
PD 20-FEB-2003.

XX 23-JUL-2002; 2002MO-EP008220.
PF
XX 23-JUL-2001; 2001EP-00117608.
PR
XX 24-MAY-2002; 2002EP-00011710.
PR
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-342400/32.
XX
PT New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.
XX
PS Disclosure; SEQ ID NO 678; 104pp; English.
XX
CC The invention relates to a novel use of irinotecan or its derivative for
CC the preparation of a pharmaceutical composition for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject having a genome with a variant allele which comprises
CC a multidrug resistance 1 (MDRI) polynucleotide. A composition of the
CC invention has cytostatic activity. The present sequence is used in the
CC exemplification of the invention.
XX
SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLMDMNVNTMTNSNPDFTKCFONTVLYWVPCFYLMACPFYFLYSRH 60
DB 1 MALRGFCADGSDPLMDMNVNTMTNSNPDFTKCFONTVLYWVPCFYLMACPFYFLYSRH 60
QY 61 DRGYIOWTPIINKTKTALGFLIMIVCWADLFYSFWERSRGIFLAEVFLVSPILLGITLLA 120
DB 61 DRGYIOWTPIINKTKTALGFLIMIVCWADLFYSFWERSRGIFLAEVFLVSPILLGITLLA 120
QY 121 TELIQLEBRKGVSSGIMLTFWLVALCALAIIKSKITMTAKEDNOVDLFPDITFYVYFS 180
DB 121 TELIQLEBRKGVSSGIMLTFWLVALCALAIIKSKITMTAKEDNOVDLFPDITFYVYFS 180
QY 181 LLLIQVLVSCFSDSPFSETIHDNPNCPSESSAFSLRITFMWITGLIVRGROPLEGSD 240
DB 181 LLLIQVLVSCFSDSPFSETIHDNPNCPSESSAFSLRITFMWITGLIVRGROPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKQPAQPKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKQPAQPKESSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFRKVLTKTGFYFLMSFFPKAIIHDLMEFSGPOLIKLLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFRKVLTKTGFYFLMSFFPKAIIHDLMEFSGPOLIKLLIFVNDTKAPD 360
QY 361 WQGFYTYVLVFTACLOTLVHOYEHI CFVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTYVLVFTACLOTLVHOYEHI CFVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRPMDLATYINMTWSAPLQYITLALYLLMLNIGPSVLAVAVMYLVNPN 480
DB 421 GEIVNLSVDAQRPMDLATYINMTWSAPLQYITLALYLLMLNIGPSVLAVAVMYLVNPN 480
QY 481 AVNMMKTKTYOVAAHMSKDNFKIKMNEILNGICIKVLKIYMWELARCKDLAROSELKYLK 540
DB 481 AVNMMKTKTYOVAAHMSKDNFKIKMNEILNGICIKVLKIYMWELARCKDLAROSELKYLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTDENNIIIDAOAPFSLALENLRPLNLTLP 600
DB 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTDENNIIIDAOAPFSLALENLRPLNLTLP 600
QY 601 MVISIVQSVSLKRLRIFLSHELBPSDIERRPVKGGTNSITVRNATFTWARSDBPT 660

Db 601 MVISSIVQASVSLKRLIFLSHELPEPDSIERRPVKGCGTNSITVFNATFTMARSDPT 660
Qy 661 LNCITFSPBEGALVAVVGVCCGKSSLSALLAEMDKEGHVAIKGSVAVYPOQAWIOND 720
Db 661 LNCITFSPBEGALVAVVGVCCGKSSLSALLAEMDKEGHVAIKGSVAVYPOQAWIOND 720
Qy 721 SLRENILFGCOLPEEPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSSGQXORVSLAR 780
Db 721 SLRENILFGCOLPEEPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSSGQXORVSLAR 780
Qy 781 AVYSNADIIYLPDDPLSAVDHVGKHIFENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840
Db 781 AVYSNADIIYLPDDPLSAVDHVGKHIFENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840
Qy 841 MSGKISBMSGYQELLARDGAPAEFLRTYASTEOBDAENGVTGSGPKKAKOMENGM 900
Db 841 MSGKISBMSGYQELLARDGAPAEFLRTYASTEOBDAENGVTGSGPKKAKOMENGM 900
Qy 901 LVYDSAGKOLOROLSSSSSGDISRHNSPTELQKAEKETWKLMEADKATGOVKL 960
Db 901 LVYDSAGKOLOROLSSSSSGDISRHNSPTELQKAEKETWKLMEADKATGOVKL 960
Qy 961 SVYWDYMKALGLFISPLSIFLMCNHVSALASNYMLSLMTDDPIVNGTQEHTKVRLSVYG 1020
Db 961 SVYWDYMKALGLFISPLSIFLMCNHVSALASNYMLSLMTDDPIVNGTQEHTKVRLSVYG 1020
Qy 1021 ALGISQGIANFGYSMAVSIIGIILASRCLHVDLHSLRSFMSFFERTPSGNLVNRFSEKL 1080
Db 1021 ALGISQGIANFGYSMAVSIIGIILASRCLHVDLHSLRSFMSFFERTPSGNLVNRFSEKL 1080
Qy 1081 DTVDMSMPEYIKMFMGSLFNVIKACIYILATPTAIIIPPLGIYFFVQRFYASSROL 1140
Db 1081 DTVDMSMPEYIKMFMGSLFNVIKACIYILATPTAIIIPPLGIYFFVQRFYASSROL 1140
Qy 1141 KRLESVRSRSPYSHFNFTLGVSVIRAFEEBERFIHOSDLKVDENOKAYPSIVANRMLA 1200
Db 1141 KRLESVRSRSPYSHFNFTLGVSVIRAFEEBERFIHOSDLKVDENOKAYPSIVANRMLA 1200
Qy 1201 VRLFCVNCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTYLNMLVRMSSEMETNIVA 1260
Db 1201 VRLFCVNCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTYLNMLVRMSSEMETNIVA 1260
Qy 1261 VERLKESETEKEAPWQIOETAPSSWPQVGRVFRNYCLRYREDLPVYLRHINVTINGG 1320
Db 1261 VERLKESETEKEAPWQIOETAPSSWPQVGRVFRNYCLRYREDLPVYLRHINVTINGG 1320
Qy 1321 EKYGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITTIIPODPVLF 1380
Db 1321 EKYGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITTIIPODPVLF 1380
Qy 1381 SSGSLRNMLDPPSOYSDEEWTSLLELAHLKDFVSALPDKLDHECAGEGENTL SVGQROLVCL 1440
Db 1381 SSGSLRNMLDPPSOYSDEEWTSLLELAHLKDFVSALPDKLDHECAGEGENTL SVGQROLVCL 1440
Qy 1441 ARALLRRTKILVDEATAVDLETFDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRIVVL 1500
Db 1441 ARALLRRTKILVDEATAVDLETFDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRIVVL 1500
Qy 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLY 1531
Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLY 1531

Search completed: December 15, 2005, 15:17:22
Job time : 188.366 secs

This Page Blank (upside)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:57:04 ; Search time 37.8285 Seconds

(without alignments)
4901.320 Million cell updates/sec

Title: US-10-665-283-6

Perfect score: 9901
Sequence: 1 MALRGFCASDGPLMDMNV.....ASVAVAKAKPFSIPDSLS 1927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 0%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7849	79.3	1531	1 DVHUR	multidrug resistan
2	4484.5	45.3	1527	2 JE0336	canalicular multir
3	3587	36.2	1545	1 S71841	multidrug resistan
4	3507.5	35.4	1541	1 S71839	canalicular multir
5	3289	33.2	1494	2 E89447	protein P57C12.4 (
6	3225	32.6	1573	2 T42119	hypothetical prote
7	3195.5	32.3	1502	2 T42216	multidrug resistan
8	2838	28.7	1515	1 S51863	cadmium resistance
9	2732	27.6	1398	2 T20434	hypothetical prote
10	2548	25.7	1478	2 T38712	ABC transporter SP
11	2384.5	24.1	1623	2 T01369	ABC transporter At
12	2356.5	23.8	1632	2 D86428	glutathione S-conj
13	2326	23.5	1495	2 E86428	probable ABC trans
14	2294.5	23.4	1144	2 T27408	hypothetical prote
15	2263.5	22.9	1559	1 S64757	probable membrane
16	2238.5	22.6	1488	2 F86428	probable ABC trans
17	2230	22.5	1516	2 F84919	glutathione-conj
18	2230	22.5	1539	2 T48059	ABC transporter-11
19	2144.5	21.7	1355	2 T00961	hypothetical prote
20	2144.5	21.7	1514	2 T52080	multi resistance p
21	2140	21.6	1515	2 T52081	MKP-like ABC trans
22	2131.5	21.5	1490	2 T47840	multi resistance p
23	2129	21.5	1545	2 T46645	hypothetical prote
24	2123	21.4	1153	2 T26883	hypothetical prote
25	2109	21.3	1545	2 T42751	sulfonamide recep
26	2104.5	21.3	1511	2 T42711	sulfonamide recep
27	2098	20.9	1546	2 T42728	sulfonamide recep
28	2069.5	20.9	1389	2 T47796	ABC transporter-11
29	2062.5	20.8	1661	2 S64800	probable membrane

30	2037	20.6	1121	2 C87973	protein Y43F8C.12
31	2014.5	20.3	1582	2 A56248	sulfonamide recep
32	2008	20.3	1592	2 S48933	probable transport
33	1995	20.1	1350	2 S68403	inward rectifier p
34	1980	20.0	330	2 JC4689	inward rectifier K
35	1923	19.4	330	2 A57616	inward rectifier K
36	1920	19.4	390	2 JC7901	inwardly-rectifyin
37	1888	19.1	1477	2 S64616	YOR1 protein - yea
38	1884.5	19.0	1037	2 T50518	ABC transporter-11
39	1805	18.2	1146	2 F84487	probable ABC trans
40	1722	17.4	946	1 JC5667	multidrug resistan
41	1693	17.1	1548	1 DV1NS	multidrug resistan
42	1687	17.0	1421	2 T34225	hypothetical prote
43	1606	16.2	1427	2 T20903	hypothetical prote
44	1575.5	15.9	1427	2 T39219	acp-binding caset
45	1558	15.7	1469	2 T50210	probable ABC trans

ALIGNMENTS

RESULT 1

DVHUR
multidrug resistance protein (cell line H69AR) - human
N:Alternate names: multidrug resistance-associated protein (MRP)
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001
C/Accession: A44231; A37495
R/Cole, S.P.C.; Bhargava, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.;
Science 258, 1650-1654, 1992
A>Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer
A/Reference number: A44231, PMID:93088080; PMID:1360704
A/Accession: A44231
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: MAPRSGTGMGRGIPATPTSPAFRTSSGCLVFTSGPV, 50-1531 <COI>
A/Cross-references: UNIPARC:UPI00001746CB; GB:L05628; NID:G1835658
A/Experimental source: Small cell lung carcinoma cell line H69AR
A/Note: Sequence extracted from NCBI backbone (NCBI:P119851); this sequence has been co
R/Cole, S.P.C.; Deeley, R.G.
Science 260, 879, 1993
A>Title: Multidrug resistance-associated protein: sequence correction.
A/Reference number: A37495; PMID:93262415; PMID:8098549
A/Accession: A37495
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-60 <CO2>
A/Cross-references: UNIPARC:UPI00001746CC; GB:L05628; NID:G1835658
A/Note: sequence extracted from NCBI backbone (NCBI:P119851)
C/Genetics:
A/Gene: GDB:MRP
A/Cross-references: GDB:136335; OMIM:158343
A/Map position: 16p13.1-16p13.1
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C/Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmem
F/661-844/Domain: ATP-binding cassette homology <ABC1>
F/678-885/Region: nucleotide-binding motif A (P-loop)
F/788-792/Region: nucleotide-binding motif B
F/1310-1503/Domain: ATP-binding cassette homology <ABC2>
F/1327-1334/Region: nucleotide-binding motif A (P-loop)
F/1450-1454/Region: nucleotide-binding motif B

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCASDGPLMDMNVNTSNPDPTKGFQNTLVWVPCFYLMACPFFLYLSRH 60
DB 1 MALRGFCASDGPLMDMNVNTSNPDPTKGFQNTLVWVPCFYLMACPFFLYLSRH 60
QY 61 DRGYIQMTPLNKKTKTALGLPLWIVCWADLFYFWSERSRGIPLAPVPLVSLTGLITTLA 120
DB 61 DRGYIQMTPLNKKTKTALGLPLWIVCWADLFYFWSERSRGIPLAPVPLVSLTGLITTLA 120

QY 121 TFLIQLERRKGVSSGIMLTFMVALVACALILRSKIMTALKEDAOVDLPFDITFYFYS 180
 Db 121 TFLIQLERRKGVSSGIMLTFMVALVACALILRSKIMTALKEDAOVDLPFDITFYFYS 180
 QY 181 LLLIQLVLSGSDSPLEFSETIHDNPPCESSASFLSITFMWITGLIVRGYRQPLBGS 240
 Db 181 LLLIQLVLSGSDSPLEFSETIHDNPPCESSASFLSITFMWITGLIVRGYRQPLBGS 240
 QY 241 LMSINKEDTSBOQVPLVKNMKCECAKTRKOPVKVYVSSKDPAPCKESSKVDABEVEAL 300
 Db 241 LMSINKEDTSBOQVPLVKNMKCECAKTRKOPVKVYVSSKDPAPCKESSKVDABEVEAL 300
 QY 301 IVKSPOKEMNPSLFKVLYKTGPFLMSFFPKALHDLMMFSGPOLIKLILFVNDTRAPD 360
 Db 301 IVKSPOKEMNPSLFKVLYKTGPFLMSFFPKALHDLMMFSGPOLIKLILFVNDTRAPD 360
 QY 361 WQGFYTVLTFVTAQOTLVHGYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
 Db 361 WQGFYTVLTFVTAQOTLVHGYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
 QY 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVILALYILMLNGPSVLGAVAMVLMPVN 480
 Db 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVILALYILMLNGPSVLGAVAMVLMPVN 480
 QY 481 AVMAKTKTYOVAMHKSNDRIKLMBILNGIKYLKJYAMELAFKDYLAIROBELKYLK 540
 Db 481 AVMAKTKTYOVAMHKSNDRIKLMBILNGIKYLKJYAMELAFKDYLAIROBELKYLK 540
 QY 541 KSAIYLSAVGTFTWCTPFLVACTPFVAVVTIDENNILDAQAPFSLALFNILRPPLNLP 600
 Db 541 KSAIYLSAVGTFTWCTPFLVACTPFVAVVTIDENNILDAQAPFSLALFNILRPPLNLP 600
 QY 601 MVISIYQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGNSTIVRAVATFMARSDPT 660
 Db 601 MVISIYQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGNSTIVRAVATFMARSDPT 660
 QY 661 LMGITFSIPEGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAMIOND 720
 Db 661 LMGITFSIPEGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAMIOND 720
 QY 721 SLRENILFGCOLLEBYRHSVIOACALLPDLILPSGDSITEGEKVNLSGGQKORVSLAR 780
 Db 721 SLRENILFGCOLLEBYRHSVIOACALLPDLILPSGDSITEGEKVNLSGGQKORVSLAR 780
 QY 781 AVYENADLYLPDDPLSAVDAHVGKHIFFENVIQPKGMLKOKRILVTHSMSTLPQVDVITV 840
 Db 781 AVYENADLYLPDDPLSAVDAHVGKHIFFENVIQPKGMLKOKRILVTHSMSTLPQVDVITV 840
 QY 841 MSGKISMSGYOEILLARDGAFABFLRTYASTEOBQDAEENGVTGVSQPGKEAKOMENG 900
 Db 841 MSGKISMSGYOEILLARDGAFABFLRTYASTEOBQDAEENGVTGVSQPGKEAKOMENG 900
 QY 901 LVTSASGQOLOROISSSSSYSGDISRHNSTAELOKAAKKEBTWKLEADKAQOTGOYKL 960
 Db 901 LVTSASGQOLOROISSSSSYSGDISRHNSTAELOKAAKKEBTWKLEADKAQOTGOYKL 960
 QY 961 SVYDYMYAKIQLFISFISIFLMCGHVASLASNWMLSLMTDDPIYNGOETHKURLSYG 1020
 Db 961 SVYDYMYAKIQLFISFISIFLMCGHVASLASNWMLSLMTDDPIYNGOETHKURLSYG 1020
 QY 1021 ALGISOGIAVFGYSMAVSIIGLILASRCALHVDLHLSILSPMSFFERTSPGNLVNFSKEL 1080
 Db 1021 ALGISOGIAVFGYSMAVSIIGLILASRCALHVDLHLSILSPMSFFERTSPGNLVNFSKEL 1080
 QY 1081 DTVDSMTPEVIMKMGSLFNVIAGCTVILLATPIAIIIPGLGIYFFVQRPYVASSQOL 1140
 Db 1081 DTVDSMTPEVIMKMGSLFNVIAGCTVILLATPIAIIIPGLGIYFFVQRPYVASSQOL 1140
 QY 1141 KRLSVSVSPYVSHNETLLGVSVIRAFEBEORFTHOSDLKVDENOKAYYSIVANRWLA 1200
 Db 1141 KRLSVSVSPYVSHNETLLGVSVIRAFEBEORFTHOSDLKVDENOKAYYSIVANRWLA 1200

QY 1201 VRLCEVNCIVLPAALFAVISRHSLSAGVLGVSYSLOQTTTYINMLVRMSSEMETNIVA 1260
 Db 1201 VRLCEVNCIVLPAALFAVISRHSLSAGVLGVSYSLOQTTTYINMLVRMSSEMETNIVA 1260
 QY 1261 VERLKEYSETEKAPMIOETAPSSWPOVGRVFRYRCIRYREDDLPVLRHINVTNGG 1320
 Db 1261 VERLKEYSETEKAPMIOETAPSSWPOVGRVFRYRCIRYREDDLPVLRHINVTNGG 1320
 QY 1321 EKVYIVGRTAGKSSLLTGLFRINESAGEHIIIDGINIAKIGLHDLRFKTIITIPQDVL 1380
 Db 1321 EKVYIVGRTAGKSSLLTGLFRINESAGEHIIIDGINIAKIGLHDLRFKTIITIPQDVL 1380
 QY 1381 SGSIKMWLDPSSQYSDDEWVTSLELAHKDFVSALPKLHCEKAGEGENTSVGQROLVCL 1440
 Db 1381 SGSIKMWLDPSSQYSDDEWVTSLELAHKDFVSALPKLHCEKAGEGENTSVGQROLVCL 1440
 QY 1441 ABALRRTKILVDEAFAAVDLETDDLIOSTIRPOFEDCTVLTATRLNTIMDTRVYL 1500
 Db 1441 ABALRRTKILVDEAFAAVDLETDDLIOSTIRPOFEDCTVLTATRLNTIMDTRVYL 1500
 QY 1501 DKGEIOEYGA PSDILOQRGLFYSMAKDAGLV 1531
 Db 1501 DKGEIOEYGA PSDILOQRGLFYSMAKDAGLV 1531

RESULT 2
 JE0336
 canalicular multispecific organic anion transporter - human
 C:Species: Homo sapiens (man)
 C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
 C:Accession: JE0336
 R:Uchiyama, T.; Hinoshita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, M.; Biochem. Biophys. Res. Commun. 252, 103-110, 1998
 A:Title: Isolation of a novel human canalicular multispecific organic anion transporter, c.
 A:Reference number: JE0336, MUID:99032812; PMID:9813153
 A:Accession: JE0336
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Repeats: 1-1527 <UCH>
 A:Cross-references: UNIPROT:O15438; UNIPARC:UP10000169859; GB:AF083552
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
 C:Keywords: ATP
 F:1306-1499/Domain: ATP-binding cassette homology <ABC2>

Query March 45.3%; Score 4484.5; DB 2; Length 1527;
 Best Local Similarity 56.6%; Pred. No. 4.6e-296;
 Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;

QY 8 SADGSDPLMDNNTVMTNSNDPTKCFONTVYLVVPCFYLMACFPFYFLYLSRRHGRYIQM 67
 Db 7 SGEIGSKFMDBNSLVHTENPDLTFCFQNSILAWPCILYLVVALPCYLLYLRHRCGYIIL 66
 QY 68 TPLNKTITAGFLMIVYCMADLFYSFWERSRGITLAVFVLVSPILGLITTLATFLIQLE 127
 Db 67 SHELKIMVIVGLVLMCVSWADLFYFSGLVHGRAPAVFVFTPLVVGVTMLATLTIQYE 126
 QY 128 RRRKVOSSGIMLTFMVALVACALILRSKIMTALKEDAOVDLPFDITFYVYFSLILQV 187
 Db 127 RLGQVOSSGVLITWELCVCAIVPFSKILILAAABGEISDPFFFTFYHFAVLVSLAL 186
 QY 188 LSCFSDSPLESETIHDNPPCESSASFLSITFMWITGLIVRGYRQPLBGSDDLMSINKE 247
 Db 187 LACFREKPPFSSAGVNDPNPPEISAGFLSLFLPFWMTKXALYGRHRLBKEKIDLSLKEE 246
 QY 248 DTSQVVPVLVKNMKCECAKTRKOPVKVYVSSKDPAPCKESSKVDABEVEALIVKSPQK 307
 Db 247 DRSQMVVQOLLLEARRKQKQETARRH-----KASAAQK---NAGEDEVLLGAPRR 294
 QY 308 EWNISLFRVLYKTGPFLMSFFPKALHDLMMFSGPOLIKLILFVNDTRAPDMQGYFT 367
 Db 295 R-KPSFLKALLATGSSFLISACFKLIQDLISFLNPOLSLIRFISNPMGSPWVGFLVA 353

QY	366	VLLPFLACQTLVHGYEHCIVSGSMRIETAVIGAVREALVYITNSARKSSVGEIVNLM	427
Db	354	GLMFCSMMQOSLILQHYHYIFVTEGKVFRTMGVYRRFALVITNSVKKASSTVGEIVNLM	413
QY	428	SVDAORFMDLAFYINNIWSPALQVILATLALMLNCPSTLACVAVVWVLMVPVNAVMAKTI	487
Db	414	SYDARFMDLAFPLMLNMAPLOIILAIYFLMONGPSTVLACVAFMVLLIPLANGAVAYMA	473
QY	468	KTYQVAHMKSKONRIKLMNEIINGIKVLVAMELAFKQVLAIRQEBELKKSAYLSA	547
Db	474	RAFQYQOMKLKOSRIKMEIINGIKVLKVAMEBSPFLKQVBCIRQGBELQILRTAAYLHT	533
QY	548	VCTFWVCTPRLVALCTPAVYVTTIDENNILDKQTAFAVSLALNIRPLNIRPLNPIVSSIV	607
Db	534	TTTFWMCSPFLVTTITLTVVYVVPNNVLDAEKAFVSVSLNIRPLPLNMLPOLISNLT	593
QY	608	QASVSLKRLRIPLSHELEPDSIERRPVMDGGTNSITVRNATPTFMARBDPPTLNGIFES	667
Db	554	QASVSLKRLQOFLSGBEILDPOSEVKRTISPG--YAITHSGITFWAODLPPTLHSLDIQ	650
QY	668	IPGALVAVVGVGGCKSLSSALLAEMDKVGEHAIKGSVAVVYVQOAMIQNDSSRENLIT	727
Db	651	VPKGALVAVVGVGGCKSLVSGALLGEMKEKLGKVMHMGSVAVVYVQOAMIQNDCTQENVL	710
QY	728	FGCQLEBPYKSVIOACALLPDLLEILPSGDRPIEIGKVNLSGGCKQKVSLARAVYSNAD	787
Db	711	FGKALNPKRYQOOLTEKALLADLEMLPGDDOTEIEGKINISGGORORVSLARAVYSAD	770
QY	768	ITLPPDPLSAVANAHGKHFEENVICPKMGLKNKTRILVTHMSXYPOVDVILVMSGGKIS	847
Db	771	ITLDDPLSAVDSHAKHIFDHVIPBEGVLAKTRVLVTHGISPLPDQTFIIVLADGOVS	830
QY	848	EMGXYOELLARDGAFAEFLRTASTEOEDDAEENGVGSGPGEKAKOMENML-----	901
Db	831	EMGPFPALQNRGSRANFLCNVAPBEDGCHLEDSTALEGADKALLIEDTLSNHTDIT	890
QY	902	---VTDAGKQLOLRQSSSSSSGSD---ISRHNSTAE-LQRAEKKEBTWKLMEAD	951
Db	891	DNDPVTYVQKQFMKQLSALSS-DGEGQORPVRPHRLGPSKVKYQVTAKKADGA--LTQBE	947
QY	952	KAQSGQVCLSVYMDWPKAIIGLFISLSTFLPMCNHVSALASVYVLSLWTDPIYNGTOEH	1011
Db	948	KAAITVLSVYMDAKAVGLCTTIAICLLVYGOSAAAIKANVWLSAATNMAADSRQNN	1007
QY	1012	TKVRLSVGALGISGGIAVFGSVASVIGGILASCLHVNDLSHLSIRPMSFFERTPSGN	1071
Db	1008	TSIRLGVYALAGILOGFLVMLAAMMAAGIOAAVLAHQAALLHMKIRSPQSFDTTBSGR	1067
QY	1072	LVNRFSEKLDVDSMIPEVIKMFMSLFPNVIAGACIVILLATPIAIIIPPLGLIIFYVOR	1131
Db	1068	ILNCRSKOIIYVDEVYLAIVILMLNSFPNALISTLVIVMASPTLPFVILLPLAVILTYLQV	1127
QY	1132	FYVASSROLKLESYSRSPVISHNETLGVSVIAPFEBQERFIHQSDIKYDENQKAYP	1191
Db	1128	FYAATSRQIKRLESYSRSPVISHFETVGSVIAVYRSDRFEIISPTKYDANDRSCYP	1187
QY	1192	SIVANRWTAIVALECGNCTIVLFAALFAVISHSLSAGLVGVSYSVSLQVTTYLMLVMS	1251
Db	1188	YIISRWLSTIGEFEGNCVVLFAALFAVIGKSLNPGVLGVSVSLSQVTTALMMIMIMM	1247
QY	1252	SEMETNIVAVERLKEVSETEKAPMOIOETAPPSMPPOGVREFPNYCLARREDIDFYLK	1311
Db	1248	SDLBSNIVAVERVKEVSKTEPTEAPVYVGSNRPEBMPRGVREFPNYSVVRIRPGDLVLR	1307
QY	1312	HINVTINGEKVGIYGRGAGKSSITLGLFRINSEABEIIIDGINIAKIGLHDIRKIT	1371
Db	1308	DLISHVHGEEKVGIYGRGAGSSMTLCLFRLEAKAKEIRIDGLNVADIGHADVRSQLT	1367
QY	1372	IIPDPVYVFGSGLRNKLDPFSQYSGEYVWTSLEIHLADPVSALBPKDKHECAGEGNTLS	1431
Db	1368	IIPDPILFSGTLRNKLDPFGYSSEBDIKWALLESHTLTTPVSSQCPAGIDFQCSBGEENTLS	1427
QY	1432	VGOROLVCALARALKTKIILVDEAYAAVDLETDDLQSTIRTOFEDCTVITIAHRLNTI	1491

```

Db      1428 VQOROLVCARLRLKSKSLVLDDEATAIIDLSTDLIQATRTFQPTCTVLTAHRLNTY 1487
|||...|||||...|||||...|||||...|||||...|||||...|||||...|||||...
Oy      1492 MDYTRVILVDKGELICEYGAPSDLLQORGIFYSMADAGI 1530
|||...|||||...|||||...|||||...|||||...|||||...|||||...
Db      1486 MDYTRVILVDKGVAFEFPSDSPANLIARGIFGYGMARDAGI 1526
|||...|||||...|||||...|||||...|||||...|||||...|||||...

RESULT 3
S71841
multidrug resistance protein, canalicular - human
C.Species: Homo sapiens (man)
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C.Accession: S71841; S71840
R.Koenig, J.; Keppler, D.
submitted to the EMBL Data Library, August 1996
A.Reference number: S71841
A.Accession: S71841
A.Molecule type: mRNA
A.Residues: 1-1545 <KOE>
A.Cross-references: UNIPARC:UPI00001746CD; EMBL:X66395; NID:g1507819; PIDN:CA65259.1;
R.Buechler, M.; Koenig, J.; Brom, M.; Kartenebeck, U.; Sprinz, H.; Horie, T.; Keppler, D
J. Biol. Chem. 271, 15091-15098, 1996
A.Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
A.Reference number: S71839; MUID:96279006; PMID:8662992
A.Accession: S71840
A.Status: nucleic acid sequence not shown
A.Molecule type: mRNA
A.Residues: 1415-1429, 'VP', 1432-1455, 'E', 1457-1545 <BUB>
A.Cross-references: UNIPARC:UPI00001746CE; EMBL:X66395
C.Genetics:
A.Gene: GDB:ABCC2; CMOAT; ABC; MRP2; CMRP; DCS
A.Cross-references: GDB:6089489; OMIM:601107
A.Map position: 10q24-10q24
C.Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C.Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
F:654-837/Domain: ATP-binding cassette homology <ABC1>
F:671-678/Region: nucleotide-binding motif A (P-loop)
F:1317-1510/Domain: ATP-binding cassette homology <ABC2>
F:1334-1341/Region: nucleotide-binding motif A (P-loop)

Query Match          36.2%; Score 3587; DB 1; Length 1545;
Best Local Similarity 47.5%; Pred. No. 5.3e-235;
Matches 746; Conservative 286; Mismatches 461; Indels 78; Gaps 20;

Oy      3 LRGCFSADGSDPLMDWMTNTS-----NDFTKCFQNTLVWVPCFYLMACFPYFL-- 55
Db      2 LEKCN-----STFWNSFLDSPEADLPICFEGLTVVMPIPLGLWILLAPQLLHV 51
Oy      56 YLSRHDRGIQTMLNKTKTLALGFLMVLCMADELFSWMESRGIFLAPVLVSFTLLGI 115
Db      52 YKSRTKSSSTTKLYALK-QVEFGFL-LILAIELALVLTESGQAYPAVAVYYTMSLY-L 108
Oy      116 TTLATFLILOERRKGOVSSGIMLT-FWLVALVCMALAIRKIMTALKEDAQVDLFRDIT 174
Db      109 GTMLIVLIIQSRCQWCQAKSWFLSEFWISILCTGFPOFLIRTLIGOD-NSNIAYSCL 167
Oy      175 FYVVFSLLLIQLVLSCEFSDRSPLESETIHDPNCPESSASFLSRITFMWITGLVRYGRQ 234
Db      168 FFIISYGFQIILIRISAFSEN-----NESSNPSSIASSITSITYEWYSIILKYKR 219
Oy      235 PLBSGDLMSLNKETSEQVPPVAVLKMKKECAKTRK-----QEVKKVYYSKD 281
Db      220 PLLIEDWEVDEENKTKTLVSKRFTHMKRELQKARRALQRROEKSSQONGARLRGLKN 279
Oy      282 PAQPKESKVDANEVEEVALIKVSPQKENPS--LFKYLVTKFGEFYFLMSFFPKAIHDMM 339
Db      280 QSQGDALVL---EDVERKKKSGTKKDVPSMWLKLKFKEFVWVLLKSFLLKLVDIFT 336
Oy      340 FSGGQILKLILKPMDTKAPWQGYFTYVLLFVAACLQTLVLHQYFHICFPGSGRIKTV 399
Db      337 FVSGQLKLILSFASDRDYLMIGYLAAILFTFALLQSFCLOCYFDCLGFKLGKVKRV 396

```

```
Qy 400 IGAVYRKAIVYMSARSSVTGEIVNLSVDAQRFMDLATTYNNIMSAFLQVILATYLM 455
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 397 MASVYKKAIVLTLNIAKREYTVGEIVNLSVDAQKIMVDMFMHMSVQIVISIFLW 456
Qy 460 LNLGSPVLAQVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 519
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 457 RELGSPVLAQVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAV 516
Qy 520 WELAFKDVLAIRQELKVLKKSAYLSAVGTFTVCTPFLVALCTFAVYVYIDENNILDA 579
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 517 WEPSFRQVQVNLARKKELKNLLAFSOLQCVAFVFPQLPVLVSVVTFSSVYLVDSNNILDA 576
Qy 580 QTAFAVSLATFNIRFPPLNIPMLVSIYQVAVSLKRLIFLSHLEPDSIERRPVQDGG 639
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 577 QKAFTSITLFTNIRFPPLSMPLFMATISSMLQASVSTERLEKYLIGDLDLTSAIRHS 634
Qy 640 GIVNSITVRNATFTMARSDPTLNGITTSIPFGALVAVVAVVQVCGKSSLSLTLAEMDVE 699
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 635 --KAMQSEASFWEHSEATVRDVLNDIMAGQVAVVAVVQVSGKSSLSLTLAEMDVEH 692
Qy 700 GHVAIKGSVAVVPOQAMIONDSLRENILFGQLEBPYRVYIQAACALPDLIELPSGDR 759
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 693 GHITIKGTAVVPOQSMIONGTINDNILFGTEFNEKRYQVLAECALLPDLIELPSGDLA 752
Qy 760 EIGEKVAVLSGGQKQVSLARAVVSNADYLPDPLSAVDAHVGKHFENYIGPKMKLN 819
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 753 EIGKKGINTLSGGQKQVSLARAVVSNADYLPDPLSAVDAHVGKHFENYIGPKMKLN 812
Qy 820 KTRLIVHMSKYLPOVAVIYMSGGKTSKMSGYOELLARDQAPAFELRTY--ASPEQEDA 878
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 813 KTRLIVHMSKYLPOVAVIYMSGGKTSKMSGYOELLARDQAPAFELRTY--ASPEQEDA 872
Qy 879 BEENVTVGVSQPKAKQW-----ENGMVYDSAGKQLOROLSSSSSSSGDISHNH 928
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 873 -----TYHDSSEERADVOGLISSVEIPEDAASITMRRENSFRITLSSSSNHRHLSL 927
Qy 929 NSTAELQAKAEKEET-----WKIMADKAOTGOYKLSVWDYMKALGFLSFLSIFLPM 983
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 928 RNSIKTRNVNLSKDEELVKGQKLIKKEFIETGKFKFSIYELQALFSLFIFLILAFV 987
Qy 984 CNHVSALASNYWLSLMT--DDEIVNGTQ--BHTKVRLSVYALGISQIANGFSGMANSI 1039
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 988 MNSVAFIIGSNLMLASMTSDSKI FNSTDYPASORDRQVAVYALGIAQGI FVFIAPHMSAF 1047
Qy 1040 GGLIASRCLAYDLHLSILSRPMSFEERTPSGNLVNRFSEKELDYDSMIPVYIKMFMSLF 1099
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1048 GFVAASNTLHKQNLNITLIRAMPREDTTPTRGIYVRFAGDSTYDDTLIPQSLKRWITCFL 1107
Qy 1100 NVIACVILIATPLAIIIPPLGLIYFVQRFYVASSRQIKRLSVSRSPVYSHFNEL 1159
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1108 GLISTLVMI CMATVFTIIVIPGLIYVSQVMFYVSTSRQLRDLDSVTRSPDIYSHFSEV 1167
Qy 1160 LGVSVITAFEEOEERFIHOSDLKVDENOKAYPSIVANWMAVRLCEVNCIVLPALFAY 1219
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1168 SGLPVIRAFEEQORFLKNEVRIDNQCVSWITSNWMLIRLELVNLTVPFSALMV 1227
Qy 1220 ISRSLSLAVGLSVSYSLQVTTYLTMVLRMSSEMETNIVAVERLKEYSETEKEAPQIQ 1279
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1228 IYRRTLSGDYGVFLSNALNITQYTNMLVVRMTSELTNIYAVERTTEYTKENBAPM-VT 1286
Qy 1280 ETAPSPSPQVQVAYEFNRYCLRYREDLDVLRHINVTINGEKVGIIVRTGAKSSLTLLG 1339
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1287 DKRPDPMPSPKGIQFNNYQVRYRPELDLVLRGICDIGSMKKIGVVERTGAKSSLTINC 1346
Qy 1340 LFRINESAGEIIIDGINIAKIGLHDLAFKTTIITQDVVLSGSLRMLMDPESQYSDDEV 1399
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1347 LFRILEAAGQIIIDGVDIASIGLHDLAKKLTIIIPQDILPESGSRMLMDPENNYSDEI 1406
Qy 1400 WTSLELAHLKDFVSLPDKLHDECAEGENSVGQROVLCCARALLRTKTLIVDEATPA 1459
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1407 WKALELAHLKSFVSLQGLSHGEGEAGGNLSIGRQLCLGRALLRKSXKILIVDEATPA 1466
Qy 1460 VDEETDILIGSTIRTOFEDCTVLTIARHLANTIMDYTRVILDKGEIQEYGAPSDLLQORG 1519
```

```
Db 1467 VDEETDILIGSTIRTOFEDCTVLTIARHLANTIMDYTRVILDKGEIQEYGAPSDLLQORG 1526
Qy 1520 LFYSMAKADGL 1530
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1527 PFYEMAKEAGI 1537
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
571839
canalicular multidrug resistance protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text, change 09-Jul-2004
C:Accession: S71839
C:Buechler, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, D.
J. Biol. Chem. 271, 15091-15098, 1996
A:Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
A:Reference number: S71839; MIMD:36279006; PMID:8662992
A:Accession: S71839
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1541 <BUE>
A:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
F:100-124/Domain: transmembrane #status predicted <TM01>
F:127-151/Domain: transmembrane #status predicted <TM02>
F:160-187/Domain: transmembrane #status predicted <TM03>
F:305-329/Domain: transmembrane #status predicted <TM04>
F:354-381/Domain: transmembrane #status predicted <TM05>
F:431-451/Domain: transmembrane #status predicted <TM06>
F:456-476/Domain: transmembrane #status predicted <TM07>
F:536-564/Domain: transmembrane #status predicted <TM08>
F:574-602/Domain: transmembrane #status predicted <TM09>
F:667-674/Region: nucleotide-binding motif A (P-loop)
F:966-994/Domain: transmembrane #status predicted <TM10>
F:1018-1046/Domain: transmembrane #status predicted <TM11>
F:1104-1133/Domain: transmembrane #status predicted <TM12>
F:1203-1228/Domain: transmembrane #status predicted <TM13>
F:1330-1337/Region: nucleotide-binding motif A (P-loop)
F:1330-1337/Region: nucleotide-binding motif A (P-loop)
F:6,1007,1010,1011/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 35.4%; Score 3507.5; DB 1; length 1541;
Best Local Similarity 46.2%; Pred. No. 1,36-229;
Matches 721; Conservative 301; Mismatches 471; Indels 69; Gaps 21;

Qy 6 FCSADGSDPLMDNMTNNTSNPDFTKCPQNTVLVWPCFYIMACPPF--YFLYLSRHDRG 63
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 FCNS-----TFWDLSIL--ESPEADLPCLFEOYVWVWIPGLFIMLAPMQLVSVRSRTRKS 58
Qy 64 YIQMTPLNKTATNGLFIMVCMADLVYSFMRERGLFLAPVPLVSPVTLGITTLATFL 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 SI--TKRYLAQVVFVLLILALIDLSLALTEDGQATVPVRVYTNP--ILVLCIWLVLVA 115
Qy 124 IQLERRKGVSSGIMLT-FMVLAVCALAIRSKIMTALKEDAOVDLFRDITFYVESL 182
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 116 VQHRQMCVNRKNSFSLFIMLSTVLCGVFOQT-LIRALLKDSNNMAYSTLFFVSVGFQ 174
Qy 183 LIQLVLSCFSDRSPFLSETHIDPNPCRESSASFLSRITFMWITGLVGRYRQPLEGDLW 242
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 IYVLLILAFSGPS-----DSYQTPSVASFLSITFSWYDRVTLKGYGHPLTLEBVM 226
Qy 243 SLNK-----EDTSQVAVLVKN-----WKKECAKTRKQPVKVVYS--SDPAPKSSK 290
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 227 DIDGFKTRVSTYSFEAAKMTDLQAKQAFQRIQKQKREKATLHGKMKQSQSDVLV 286
Qy 291 VDAREEVALIVKSPQKEMNPSLFVLYKTFGQPFVLSFPFKAIHDLMMFSGPOLIKLLI 350
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 287 LEBAKKSEKTKYQYPSW--LKSLSFKTHVIVLSFIKLHDLVLFNPLQILKLI 343
Qy 351 KFNVDTRAPDWQGYFYVLLFVTACIQTLVLHGYFHIQVSGMRIKTAVIGAVYRKALVI 410
```

```

Db      344 GFVSSNSYVWFYICAILMFAVTLIQSFCLQSYFOHCFVLGMCVRTVWSSYKKALLTL 403
Qy      411 TNSRKSTVEIYVWMSVDAQRFMDLATYINMTWSAPLOYITLALYMLTSGSVLAGV 470
Db      404 SNLARKYTTIGETVWMSVDSQKLMDATNTWQVWSSVIOITLISIFLMBELGSSILAGV 463
Qy      471 AVWLVAVVAVMAAMKTKTYOVAHMKSKDNRIKLMNEITNGIKYUKUYAMELAFKDYLA 530
Db      464 GWVLLIPVNGVLTAKIRNIOVQMKDKRKLIMNELSGIKLTKFAMEPSPFOBOVOG 523
Qy      531 IROBELVULKSAVLSAVGFTWVCTPELVALCTPAVYVITDENNITLDAQTAFFVSLAFN 590
Db      524 IRKXELKNLRFQGLQSLIFILQITPLVSVTFVSYVAVDSANVLAEMAFSITLLEN 583
Qy      591 ILRPLNLTWVVISIYQASVSLKRLIFLSHBELEPDSIRRPVKOGGNSITVRNAT 650
Db      584 ILRPLNLTWVVISIYQASVSLKRLIFLSHBELEPDSIRRPVKOGGNSITVRNAT 650
Qy      651 FTWASDPPTLNGITFTSIPGALVAVVGVGCGKSSILSLALAMDYVEGHVALKGSVAY 710
Db      640 FTWDPDEBATIQQVNDLIRKQOLAVVGVGSGKSSIVASMLGEMENVHGHITIQGSTAY 699
Qy      711 VPOQAWIQNDSLRNLTFCQLEBPYRSVIQACALLPDLIELSPSGDTEIGEKVNLG 770
Db      700 VPOQSWIQNGTIKONILFGSEYNEKTYQVLAACALLPDLIELFGDMAEIGEKVNLG 759
Qy      771 GOKORVSLAAVYSNADITYLEDDPLSAVDHVGKHFENVIGPYKMLKNKTRILVTHSMS 830
Db      760 GOKORVSLAAVYODADIYILDDPLSAVDHVGKHFENVIGPYKMLKNKTRILVTHGHI 819
Qy      831 YLPQVNDIYVMSGKSIEMGSYOEELARDGAFAELFTYAS--TQEOBDAENGVYVSGP 889
Db      820 FLTPVDIEIVLGGKTIIEKGSYRDLDDKGVFANMKTFTMGSGPEEBATVNN----- 872
Qy      890 GKEAKQWENQMLVT-----DSAG-----KOLQROLSSSSSYSGDISRHNNSTAEIQ-- 935
Db      873 DSELEDDEDDGLIPTMEIIPEDASLAMBRENSLRITSSRSSRSRSGKSLAKSLTKVY 932
Qy      936 ---KAEAKKEETWKLMEADRAQTQGVLSVYWDYMKALGLEISFLSIFLFWCNHVASLAS 992
Db      933 NVLKEKEVEGQGLIKEFEVETGKVFSTYKLYLQAVGWMSILFILLFGANNVARTGS 992
Qy      993 NYVLSLMT--DDPIVNGT---QEHKXVLSYVGAIGIOGIAVREYSNAVSGIGLARCL 1048
Db      993 NMLWLSAITSODNLTNGTSSSHDMKRIGVFAGLQAGICLLISTMSIYVACRNASKAL 1052
Qy      1049 HVDLHSLRSPMSFFERTPSGNLVNRFSEKELTVDSMIPEVIMFGSLFENVIGACIYI 1108
Db      1053 HGQLLTNITLAPRPFPTTTPRGIVNRFSGDISIVDDLLPQTLRSMWMCFFGIAGTLYMI 1112
Qy      1109 LLAETPLAIIIPPLGLIYFVQRFYVASSRQLKLESVSSPYVSHNETLLGVSVIRAF 1168
Db      1113 CMATPVPALIIIPLSILYISVQVYVATSRQLRDLDSVTKSPISHSSEVTVGLPIRAF 1172
Qy      1169 EBOERFIHQDLKVDENQKAYPSIYANRMLAVRLECVGNCIYLFALLPAVISHSISAG 1228
Db      1173 EHOORFLAWEKQIDINQKCVFSWITSNRMLAIRLELVGNDLVVFCALLVIYRKILTGD 1232
Qy      1229 LVNGLSVYSLOTVTTYLWMLVRMSSEMETNTVAVERLKEYSETEKAPMOIQETAPSWP 1288
Db      1233 VGVFVLSNALNITQTLNMLVVRMSEMETNTVAVERISEYINVENEAEM--VTDKRPDPWP 1291
Qy      1289 QVGRVEFRNYCLARYEDLDVLRHINVTINGEKVGIIVGTGAGKSLTGLFRLINESAE 1348
Db      1292 RHGEIQNNYQVRRPELDVLRKGTGNTGKSGEKVGVGRTGAGKSLTGLFRLINESAE 1351
Qy      1349 GEIIIGCINIAKIGHLRKRTIIPDDPLVFGSSGLFMNIDPVSQYDEDEVWMSLELAH 1408
Db      1352 GQIILIDIVASIGLHDLRELTIIIPDDPLVFGSSGLFMNIDPVSQYDEDEVWMSLELAH 1411
Qy      1409 KDFVSALPDKLDBHCAEGGENTLVGQROLVCLARALARTKILVLDATAVADLETDDL 1468

```

```

Db      1412 RSFVSGQLGLSHEVTEGDNLSIGRQLCLGRAVLRSKIIVLDEATAVADLETDSL 1471
Qy      1469 QSTRQGFEDCTVLTIRHRLNTMDYTRVVLVDGELQEGABEDLLQORGLFSPMAKDA 1528
Db      1472 QTIIRKESQCTVTTIARHLTIIMDSKIMVLNKGKIVEGSPBELLNSRGSFYLAKEA 1531
Qy      1529 GL 1530
Db      1532 GI 1533

RESULT 5
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: E89447
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A/Reference number: A75000; MUID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.el
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A/Accession: E89447
A/Status: preliminary
A/Molecule type: DNA
A/Releaser: 1-1494 <STO>
A/Cross-references: UNIPROT:Q20943; UNIPARC:UPI000017801E; GB:chr_X; PIDN:AAA63299.1; P
C/Genetics:
A/Genes: F57C12.4
A/Map position: X
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match      33.2%; Score 3289; DB 2; Length 1494;
Best Local Similarity 47.6%; Pred. No. 9.3e-215;
Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;

Qy      119 LAFLIQLERKRGVQSSGIMLTWVLAVALCAALMSKINTALKEDAQVDLPFDITTYVY 178
Db      76 LAFLIYVACNKKKIIITSGVITLWLVVVGIGIEFRFYLSGFIYNEVLAIGIRATLYTIA 135
Qy      179 FSLILQLVLSGCSDSRPLSEITIHDPNCPRESSASFLSITPMWTGLVLRGYRPLEG 238
Db      136 FTSSALELFLCCPAD--VSDMYKSSSCPBTTASTINLTITQWFTGLVYLGKSLLEN 192
Qy      239 SDLMSINKEDTSBOVVPVLVKNWKKCAKTRKQPVKVYVSSKDP--AOPRESSKVDAEE 296
Db      193 EDLMDNEIDKANKLIPSFVQNLKPRIDEVH-QNIK-----KOPSAALPKN----- 237
Qy      297 VEAIIYKSPQKKNPPLSFKVLVYTFGPFYLMSPFFRAIHDLMFSGPQIILKILKYNDT 356
Db      238 -----HPSFVLPPIFKTYKYTLLAGFPFYKLCFMDQFLAPQLIKOLIGFTEDK 284
Qy      357 KAPDMQGFYTVLTPYACQTLVLRHOYFHICVSGMRITAVYIGAVYRKALVITNSARK 416
Db      285 NQVWMTGCSIVGIMFSSFLQSFPLHQYHSMRILGHRVSVTSAYTSRALNLSNARK 344
Qy      417 SSTVGEIVNLSVDAORFMDLATYINMWSAPLOVILATLYLNLNGPSVLAGVAVVLM 476
Db      345 GKTIQAIWNMSVDIQKIDMAPMTLFWSPAPQIFLSTIFLWKFPLGVALAGVVLILA 404
Qy      477 VPVNAVMAAMKTKTYOVAHMKSKDNRIKLMNEITNGIKYUKUYAMELAFKDYLAIRQEL 536
Db      405 LPVNGILAIQMRKQTEOMKLDKERTKMSSEILNKGKVLKLYSWERSMEMVYKIREREL 464
Qy      537 KYVKSAYVLSAVGTFTWVCPPELVALCTPAVYVITD-ENNIIIDAQTAFFVSLAFNIRP 595
Db      465 HILKULSTFMAAVFVSCAPFLASVTSFVYVYILPENNVLTPETLFAALSFDILKRP 524
Qy      596 LNIIPVVISIYQASVSLKRLIFLSHBELEPDSIRRPVKOGGNSITVRNATFTWAS 655
Db      525 LAVAVAVVYGAVQGSVSNTRLKEFFAAEEMSPOT---SISHGETDAIEVENGLPSWSS 580

```

Qy	656	DDPPLNGTISIPEGALVAVVGVCCKSSLSLSLLAEMDVVEGHVALKGSVAVYPOQ	715
Db	581	DEDPTRISFTIQGOLVAIVGKSGSSLSLHLGEMNKLSSGVQNGINIAVYPOQ	640
Qy	716	WIQNDLSLENTLFCQOLEEPYRSYIOACALLPDLIELPSGRTIEGKVMISGGQOKR	775
Db	641	WIQNWSLNNLILFNKPYDLENEDVYKCALKEDLANLPAGRTIELGEGINILSSGQOKR	700
Qy	776	VSLARAVSNADIVLPDPLSAVDHAVHKGHIFENVIIGPK-GMLKNKTRILVTHSMSYLPQ	834
Db	701	VSLARAVYQNPDIILLDDPLSAVDSHVKGHIFENVISSSTGCGLASTRYLVATHGLTYLKH	760
Qy	835	VDVLIYMSGCKTISEMGSYOELLAROGAFPELRTASTBOBDAENGYTVSGRC----	890
Db	761	CDQLVILKEGTISELGTYOELLNNSGAFPELFEELIESKTRGRVAYSIGDSSGEVDEIL	820
Qy	891	KEAKOMENGMLVTDSAGKOLOROLSSS-----SYSGDISRH-----HNSTRELQK	936
Db	821	RDLOGVKRGIL-----KRLBSHLQOESPKEDTSARATISYSDSSSKRSVLHMSPSQCHE	874
Qy	937	AEA-----KKEETWKLMEADKAAQTGOVLSTYWDYMKALGLEFISPLSIFLFCNHV	987
Db	875	NEALIGALISEDVPQAOENTOLIEKETVETGKVFEEVIAFYQAIMSIPITLTFEFLYVSSGS	934
Qy	988	SALASNYWLSMTD-D-PIVNGQOHTKRYLSYYGALGISQGIAPFYSMAVSIIGLILAR	104
Db	935	LGISLSNFFLAKLSDAKSGNRTSSDAKMLGTYAVYWGQOSFVVLASTLITLTGLVIRAR	994
Qy	1047	CLAHVDLHSILRSPMSPFERPTSGNMLVNRSPKELDPVDSMIPEVIMKMFSSLFENVIGACI	1100
Db	995	ILHGLGLENIRKSPMAFPDVTPIGILNIRIGDIEADIDTLDVIRHMSMTIFENVATIV	105
Qy	1107	VILLATPIAAITIPGLGIYFPVQRPYYAASSRQKRLBSVSSSPVYSHNETILGVSVIR	116
Db	1055	VIMATPMAGIAFALLISVIYFVLTFYISTSRQKRLBSASRSPISYSHFOESIQGASSIR	111
Qy	1167	APEBERFIHOSDLVDENQOKYVYSIVANRMLAVLBEVCNGICYLFAALPAVISHS--	122
Db	1115	AFGVVDNFTKOSQOQRVDHLLAYYYSIVANRMLAVLLEWNGMLIVLSAAGAAVYFPDSDG	117
Qy	1225	LSAGLVGVSYSXLQVTTVYLANLVMSSEMENTNIVAVERLKESETEKEAPMOIOETAPP	128
Db	1175	LSAGLVGVSYSXALNI TOTLANAVYMTSLEFINIYSVEIKEXYTVPTGG--NNSRLAA	123
Qy	1285	SSWPQGVKEPFRNYCLRYREDLDVYLRIHINTVINGEKVIGIRTGAGKSSLTGLERIN	134
Db	1233	KSWPEKGEISIKNFVRYRPGDLVYHGISAHIAPSEKGIQVGTAGKSSLTALFRIT	129
Qy	1345	ESABEELIINDGINIAKIGLHDLRFKITTIPOBPVYFSSGLRMLNLPFSQSDSEVWTSLE	140
Db	1293	EADGGSIEIDGINIANLQLEQJRSCLTIVPODPVLFSGTKRNLDPFSAYSOSQWEALE	135
Qy	1405	LAHLDPFSAJLBDKLDEHCEAGGSENL-----SYGOROLVCLARALJLAKTKI	145
Db	1353	NAHLKPFYKSLQDGLFHKISBEGENLRPAKSSQKQLNPFVSQOROLICLARLLRKTYV	141
Qy	1451	LVLDEATAAVDLETDDLIOSTIRTOFEDCTVTITIAHRLNTIMDYTRVILDKGEIOEYGA	151
Db	1413	LVLDEAATAAVDETSSLIOKTRIREQKECTVTITIAHRLNTVWDSRLVLVDKGRVAEPDS	147
Qy	1511	PSDLL-QORGLFYMAKDAGLV 1531	
Db	1473	PKNLLANPDGIFYMAKDAANV 1494	

RESULT 6

T21219
hypothetical protein F21G4.2 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C:/Accession: T21219; T24002
C:/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

R;Mortimore, B.
submitted to the EMBL Data Library, October 1996

[illegible]

QY 668 IPEGIVANVQGVGGKSLSLALLAEMDKYBEGHAIIGSAVYVQQAQMIQNDLSRENTL 727
Db 665 VNRGGLVTVVGVGKSSMLQALMGEMEKLSGISIMHGRLCYVYQQAQMIQNDLSRENTL 744
QY 728 FGCQLEPEYRSVIVQACALLPDLEILPSGDRTEIEKQGNLSGGQKQVSLARAVYNSAD 787
Db 745 FKGQDEYFYSVNLADACALYRDLQILLPGDNIIEKQGNLSGGQKQVSLARAVYNSAD 804
QY 788 IYLFDDPLSAVDAAHAKHIFENVIGPKKMLKNKTRILVTHSKNSYLPQVDVITVNSGKIS 847
Db 805 IYLLDDPMSAVDAHAHQSLFGSVIGPEGMRLNKTRILVTHNELSFEKSDLLIVNMEGKIE 864
QY 848 EMGSYQELLARQAFAEFL-----RYASTQEQDQDAENGVTVGSGGKAKQKQEN 898
Db 865 YSGKTYDDLM-QQGAFFBOLLICEKEKEREKEASADDEDSSEFGIMIGDSDFEYDD 923
QY 899 GMLVTVDSAGKQQLORQLSSS--SSYSGDISRHNHSTA---ELQKAPAKEETWKLMEA--- 950
Db 924 DYM-----ASPIIDHVLGTHSHSTVSGIINRRRLISTHKKQRRLSTTSHHSTHSATQ 979
QY 951 -----DKAQTQGVKLSVYMDYKAIQLFISFLSIFLPMCHVVALASNYWLSMTDDP 1003
Db 980 TRQLTGTRVEVERGVKMDPTYYKYFGAMGMSIAVLVFLGMWTSTIFSMGRNMLTDMWSDN 1039
QY 1004 IV-----NQTQHTKVRLSVYGLGISQGIAGVGYMAVSGIGILLASRCLHVDLHSLIRS 1059
Db 1040 AARSSNSTGQIARILRGVYAGLGFSEIILLFTGMLSLIGGVSSRNLHAPLMKNLFRV 1099
QY 1060 PMSFPEPTPSGNLVNRFSEKELDTVDSMTPEVIMKMGSLFNVNIGACIYILATPAAITII 1119
Db 1100 PMAFVDTTPFGILNRIGKDIETVDVLLPENVQFPAQCLLVVSTLIIIMISTVPGVIV 1159
QY 1120 PPLGLIYFFVQRFYVASSRQLKRLSVSRSPYVSHFNETLLGVSVITAFEEQERPHQSD 1179
Db 1160 IPLSVYMLVMWRKYVYATSRQLKRLSITRSPYSHLSIESIGSATIRAYHLYDRFCRLSE 1219
QY 1180 LKVDNOKAAYPSYIVANMLAVRLCEVNCIYLPALPAVLISRSLASGLVGLSVSYSLQ 1239
Db 1220 TKVDSHVQCRYIVANMLSVRLPFTGNCIYLPALPAVLIRRTTISGVIGLSVSYSLN 1279
QY 1240 VTTYLVNMLVRMSSEMETNIVAVERLKEYSETKEAPMOIQE-TAPSSWPQVGVREFRNY 1298
Db 1280 ITTVLVNFAVROQTUKLETNIVSVERYKEYAEFETAEKWSSEPKQWPEBGHAIYNNY 1339
QY 1299 CLRVRBEDDFVLKRVNVTINGEKKYIGVRGAGKSSITLGLFRINSEABEGIIIDGINI 1358
Db 1340 SARVPFGMLVYKQLVNVEIKHEKVGIVRGAGKSSYTLSLFRITAEABEQIIVDGINTL 1399
QY 1359 AKIGIHLDRPKFTIIPPOPVLPFSGSLRMNLPFPGSYGEEVWTSLELAHLADFPASLPDK 1418
Db 1400 AEIGIHLDRSNITIIIPQVPVLPFSGTLRPNLDPFNYSISGDITWKTLEMMNLKEPFAHNEQ 1459
QY 1419 LDHECAEGEENISVGQRLVCLARALLKRTKILVDEKTAADVLETDLLIOSTIRTOFED 1478
Db 1460 LNYITTEGGDNISVGQRLVCLARALLKRTKRLILDEKTAADVSTOALLIKTIREEPAN 1519
QY 1479 CTVLTIARLNTIMDYTVIYLDKGEIOEYGAPSDLOQRL-FYSNAKDNGLV 1531
Db 1520 ATVLTIARLNTIMDYRIIVLNDGKVEFPSPANLISNRNSEFYSNAKRAGLRI 1573

RESULT 7
T42216
multidrug resistance-associated protein homology MRP-1 - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T42216
R.H.Itohahshi, T.; Suzuki, H.; Ito, K.; Ogawa, K.; Kume, K.; Shimizu, T.; Sugiyama, Y.
Mol. Pharmacol. 53, 1068-1075, 1998
A>Title: Hepatic expression of multidrug resistance-associated protein-like proteins mat
A/Reference number: Z22081; MUID: 98279126; PMID: 9614210
A/Accession: T42216

A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-1502 <HR>
A: Cross-references: UNIPROT:O88269; UNIPARC:UPI000012F57A; EMBL:AB010466; NID:G3242457;
A: Experimental source: strain Sprague-Dawley; liver
C: Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 32.3%; Score 3195.5; DB 2; Length 1502;
Best Local Similarity 43.1%; Pred. No. 2.1e-208;
Matches 655; Conservative 290; Mismatches 525; Indels 51; Gaps 9;

22 WNTSNPD-----FTKCFONTYLVWPCFYLMACPFYFLYSRDRGTYOMTPIKTKT 75
Db WNTQGEVVAHNLMLCFILRAAGSWPMPMTLVLCPIYLLIHRGCTLRKSRILFKIKM 79
76 ALGFLMLIVCWADLFYSFWERSRGIFLAPVFLVSPFTLLGTTLLATFLIOLERRKGVSS 135
Db VLGFRLILLYTFNMAVPLRIHRGMPQAPPELLIHTTWLTMSPATFLIHERKKGVAS 139
136 GIMTLFWLVALVALAILRSKIMTALIKEDAQVDFRDTTFYVYFSLLIQLVLSCFSPDRS 195
Db GLIFGYW---LLCCVLPAIDTVQOASAGSFQEPFLHILATYLCLSLVAELVLSGLVDQ 196
196 PLFSTHIDNPCEPSSASPLSRITFWWITGLIYRGYQPLEGSLNLMNEDTSEQVYP 255
197 PFFSDSKPLNCPPEAEASFPFKAFWASGLLWGRKYLGLPKDLMSELERNSSSELV 256
256 VLKWKMKCAKTRKQPVKVVYSSKDPAPQKSSKVDANEEVEALIVSPKEMNPSLFK 315
Db QLEKMRKN-----FSELPGHKHSGM--GTPEFALV--QPEKSGQPLIR 299
316 VLYKTFGYPLMSPFFKAIHDLMFSGPQILIKLKFVNDTKAPDMOGFYTVLLFVTAC 375
Db AIMRVFRTSTPLGLTLSTLVSDAPFAVPKLLSLFLEFGDLESSAMTGMLAVLMEISAC 359
376 LQTLVLDHYFHI CFVSGKRITAVI GAVYRKALVITNSARKSSTYGEIVNLSVDQKPM 435
360 LQTFEEQYMWKVLQWRKRLTAITGLVYRKVLVYSSGSRKSSAAGDVNVLVSVDQRLV 419
436 PLATYINWASAPLOVTLALYLLMLNLTGSPVLAGVAVWLVVAVVAVAMTKTKYOVAMH 495
Db ESILHNLMLFLMFLIIVCFYVYLMQLGSPALTAIVAVLSLPLPLFTIKRSPFOEOM 479
496 KSKNRIKLAMEIINGIKVLKYAMELAFKQKVALIROBELKVLKKSAYLSAVGTFTWC 555
Db RQKSRARLTSSMLRTVATIKSHGMECAFRLRLIHRQELGALKTSFLPSVSLVSPQ 539
556 TPFLVALCTPAVYVYTIDENNLIDQTAFAVSLAFNILEPFLNILEPMVYSSIVQASVSLK 615
Db STFLVALVFAVHTLVIEDNAMDEKAFVTLVTLVSIINKAOFLPFSVHCAVQARVSPDR 599
616 LRIFLSHELEPDSIERRPVMDGGTNSITVRNNTFTMARBDPPTLNTISIPGALVA 675
Db LAFLFCLSEEDVPNGVNSPSR-CSSKQDISIHNGFPANSSQSPCLHINILTVPGCLIA 658
676 VVGQVGCSSILSATLAEMDKVEGHVAIKGSVAVPOQAIQNDSTLENTLIFGQLEEP 735
Db VVGPRGAKSSLSLALGELLKVEGVSISIESSVAIVPEAVYQNTSVENVCFRBELDP 718
736 YRSVYQACALLPLEILPSGDRTEIGEGVNLSSGGQKQKRVSLARAVYSNADILYFDPL 795
Db WLQEVLEKCALGSDVAFSPAGVHTPVVGQGNLISGQKQKRLSLARAVYRAAVYLMDDPL 778
796 SAUVAHVGKHLFENVIGPKKMLKXKTRILYTHMSYLPQVDVILYMSGGKISEMKSQYEL 855
Db AALDAHVSQEVFKOVITGPGSLGQGTIRLVTHTLVLPQADQILVLAVAGTTAEMGSYD 838
856 LARDGAFAFLRTASTEOQDAENGVT-----GVSGPKKAKOMENGLVTDLSAQK 909
Db LHRNGALVGLDGAHQAPAGEGEGAHAAATSDLDGFGCGGTPIRRRP----- 886
910 LQRLQSSSSSYSGDISRRHNSTAELOKAEAKKETWKLMEADKQQTGOVYKLSVYWDYKA 969

Db 887 -RRPSPDAAPVKS-----TSEAQMESPSSLDDEVTVTGLTAGEDSVQYGRVKSATYLSLIRA 940
QY 970 IGLFISFLSIFLPMCNHVSALASNYMLSLMTDDEIVNGTOEHYKRLSVYALGISOGIA 1029
Db 941 VGRPLCYTTLFLFLCQGVASFCQGVWISLMDDEVVNGKQHSALRSGISFBLGLCLQAIG 1000
QY 1030 VFGYSMAVSTGGIILASRCLHVDLHSLRSWMSFPERTPSGNLVNRSKELDTYDSMTP 1089
Db 1001 LFSMAVAVFLGASASCLFLRSLMDVARSPIGFEEPTPVNGLNRSFKETDIDVDVIPD 1060
QY 1090 VIKMFMSSLFNVIGACIVIIATPLIAITIPGLGIYFVGRFVYASSRQIKRESYSRS 1149
Db 1061 KMRLLTYARGLVGLVAVSAPATLVAIILPLMLTYAGFQSLTYATCCQLRLESASYS 1120
QY 1150 PVYSHFNETLLGVSVIRAFERFERFIHOSDKVDENOKAVYPSIVANRWLAVRLECVGNC 1209
Db 1121 SVCHLAETFGQGVAVAFQAGFPTAHDALMBENORISPRVLADRWLAAANIELLNG 1180
QY 1210 TVFPALEPAVISRHSLSAGVLGSVSYSLQVTTYLNLVMSSEMETNIVAVERLKEYSE 1269
Db 1181 LVFPAATCAVALSKAHLISGLAGFSVSALQVTQLQWVVRSMVDLENSMVAVERQDVH 1240
QY 1270 TEKAPQIOETAPSSSPQVGRVFRNYCIRYREDLDFVLRHINVTNGEKYGIYRT 1329
Db 1241 TPKEAPRPLPSSAAQPLMPCGGQLEFRDPRHRPELPMAYQVSLKIHAGEKXGIYRT 1300
QY 1330 GAGKSSITLGLFRINESAEGIIIDGINIAKIGLHDLRFKITIIPDFVLPSSGLRMULD 1389
Db 1301 GAGSSSLTWGLRLQEARTEGIMWDGVPTMGHLTSLRSTITTIIPQDPVLPSSLRMULD 1360
QY 1390 PFGYSDEEVTSLLEHLKDFVSALPDKLDECAEGENLSVQGRQLVCLARALLRRTK 1449
Db 1361 LLAGNTDEGIIWALETQKAFVTSLPQOLGYECGGDDLSVCGQLCLARALLRRTK 1420
QY 1450 ILVDETAAPADELTDLLIOGTIRTOFEDCTVLTIAHLNLTIMDTYIVDKXEIOEYG 1509
Db 1421 ILIIDEATVAVSDPGETIOMQALERWFAQCTVLLIAHLRSVMCAVLMDEQVASEG 1480
QY 1510 APSDLQORGLFYGMAXDAGL 1530
Db 1481 SPAQLLAQKGLFYRLAQSGL 1501

RESULT 8
551863
cadmium resistance protein YCF1 - Yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YD9302.11c; protein YDR135c
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S51863; A55352; S50233
R:Oliver, K.; Harris, D.
submitted to the EMBL Data Library, February 1995
A:Reference number: S51863
A:Accession: S51863
A:Molecule type: DNA
A:Residues: 1-1515 <OLI>
J. Biol. Chem. 269, 22853-22857, 1994
A:Title: A yeast metal resistance protein similar to human cystic fibrosis transmembrane
A:Reference number: A55352; MUID:94357936; PMID:7521334
A:Accession: A55352
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-679, 'R', 681-1515 <SGZ>
A:Cross-references: UNIPARC:PI000013A67B; EMBL:248179; NID:g665657; PID:
C:Genetics:
A:Gene: SGD:YCF1; MIPS:YDR135c
A:Cross-references: SGD:S0002542; MIPS:YDR135c
A:Map position: 4R
C:function:
A:Description: required for cadmium resistance
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein; yeast vacuole
F:287-308/Domain: transmembrane #status predicted <TM1>
F:345-366/Domain: transmembrane #status predicted <TM2>
F:421-442/Domain: transmembrane #status predicted <TM3>
F:446-467/Domain: transmembrane #status predicted <TM4>
F:534-555/Domain: transmembrane #status predicted <TM5>
F:558-580/Domain: transmembrane #status predicted <TM6>
F:646-829/Domain: ATP-binding cassette homology <ABC1>
F:663-670/Region: nucleotide-binding motif A (P-loop)
F:951-972/Domain: transmembrane #status predicted <TM7>
F:995-1016/Domain: transmembrane #status predicted <TM8>
F:1068-1088/Domain: transmembrane #status predicted <TM9>
F:1092-1113/Domain: transmembrane #status predicted <TM10>
F:1179-1200/Domain: transmembrane #status predicted <TM11>
F:1208-1229/Domain: transmembrane #status predicted <TM12>
F:1289-1483/Domain: ATP-binding cassette homology <ABC2>
F:1306-1313/Region: nucleotide-binding motif A (P-loop)

Query Match 28.7%; Score 2838; DB 1; Length 1515;
Best Local Similarity 40.3%; Pred. No. 4; se-184;
Matches 632; Conservative 295; Mismatches 495; Indels 146; Gaps 27;

28 DFTKPCNTLVWPCRYLACPFYF-----LVLSRD-----RGYIOMPLNKTKT 75
Db 27 DFTQCFIDGVYL-----NLSAIFMTITGIRDVNLCKKXSGIKYRNWITVSMALVYL 81
QY 76 ALGFLMIVCWADLFYSFWRSGIFLAPVFLVPTLLGTTTLATPLQLERKGYOSS 135
Db 82 EIAF-----VSLASINISKEAEN---PTIVSQYASTWLSLFLVALAHMIEYDR--SVVAN 132
QY 136 GIMTFPLVALVCALILRSKIMTALKEDAQVDLFRDITTYVYFS-----LLLIQLVLS 190
Db 133 TVLLFYWLFETFGNFAKLINILIRHYEG-----IMVSGQTFILTLFQVITC 180
QY 191 FS-----DRSP-----FSETHDPNCPBSSASLRLITFMWIGLIVRGYRQLE 238
Db 181 ASILLLEALPKPKPLPHQIHQITLRKKNPYDSANIFSTTWSMGLMKTGYEKLVE 240
QY 239 SDLWSLNKEDTSEQVNVFLVGNMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVE 298
Db 241 ADLYKLEPRNFSSELSQLEKRMENL-----KQKS----- 271
QY 299 ALIVKSPQKWNBSLFVYLKTFGPPFLMSFFPKAHDLMFSGPOLIKLIRKVN--- 355
Db 272 -----NPSLSWAI CRTFGSKMLLAFFKAIDVLAFTQPQLRLILIKFVTDYNS 320
QY 356 -----TXAPDWQGYRYTLLFPTACQLVLVHQYHICVSGMRKTKA 398
Db 321 ERDDHSLQGFENNHPQKLPYRAGFLIAPAMFLVGFQTSVLHQYFLNVFNTGMYIKSA 380
QY 399 VIGAVRKALVITNSARKSSTVGEIYVNLMSVDAQRPMDLATYIMWISAPLQVITLALYVL 458
Db 381 LTLALTYQKSLVLSNEAGSLSTGDIYVLMSSVDYQKLODLQWMLNLSGPFQIITICLSL 440
QY 459 WLNAGPSVLGAVVAVMLVPPVNAVAMKTKTYQVAHNSKDNRIKLNELINGIKVLKLY 518
Db 441 YKLGLSNMVGVIILVIMPLNLSFIAMIKQKLGQSKQKXDERFRVISEIILNIKSLKLY 500
QY 519 AMELAPDKVLATLQSELKYLKKSAYLSAVGTWTWCPTPLVALCTPAAVYVITDENNIL 577
Db 501 AMERPYREKKEEVNKNELKNTLQCYMVVTSQFNIIVPLVSCFFPAVAV--VTEBRAL 559
QY 578 DAQTAFLALFNLRLPPLNLTLPVVISIYOASVSLRLRLFSEHELEPDSERRPVKD 637
Db 560 TTDLVFPAITLFINLSPLMIIIPVNLNSFLBASISIGRLFFPNEBLQPSVORLEKYK 619
QY 638 GGGTNSITV--RNATFTWASD--PPTLNGITFSIPBGLVAVWQVCGKSSLSALLAE 694
Db 620 NIGGVALINIGDADFIMQKREYKVALKNINFQAKKNLCTIVGKVSCKALLSCKMLGD 679
QY 695 MDKVEGHALVIGSVAYVPOQAWIONDSLRENILFGCQLEEPYTRSVIQAALLPDLBTLF 754
Db 680 LFRVKGATVAGSVAAVSVQVPMWNGTVKENILFGHRYDAEFYBKTIKACALTIDLAITM 739

[illegible]

C:Genetics: A:Gene: CESP:E03G2.2
A:Map position: X
A:Introns: 73/3; 113/3; 176/1; 216/1; 258/2; 384/3; 470/3; 549/1; 644/3; 757/3; 764/2;
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 27.6%; Score 2732; DB 2; Length 1398;
Best Local Similarity 41.1%; Pred. No.6.4e-177;
Matches 591; Conservative 287; Mismatches 472; Indels 88; Gaps 19;

QY 115 ITTLTATLPIQLERKGVSSGIMLTFWLVLVCALALRSKIMALKED--AQLDPRD 172
DB 27 LALLIATNEV---RRAGIHSSGPLEFCIMMLPAVNAVPEFYQMTTSGSPELVARIDFRRY 83
QY 173 ITPVYFSLIILQLVLCFSRSPLEFSETHDPN----PCPSSASFLSRITFWITGL 227
DB 84 VAVLYFPLVAEFVLHFVSDPFPM-----PRGYQMLKCPENANFISRLILMFQI 136
QY 228 IVRGRQQLBESDLMSLNKEDTSEQVYFVLVKNMKBECAKTRKQVYK-----VYSS 279
DB 137 ISLGERLTVADDFEMDSQMDQELKAKMTTEMLKQTEKAREKQVKLDDKRRERATGSE 196
QY 280 KDPAQPKKSKVDAN--EEVEALIVKSPQKEMNPSSLFKVLVYTCGPFYFMSFPFKAIHDM 338
DB 197 KAPLLGTNNYGAVALDDKRYIVQ-----PSVITLMQIMKEILIGSFIFKLSDL 249
QY 339 MFSGQILKLILKFNVDTKAPDMOGFYTVLLFVTACLOTIVLHOFHICFVSGNRIKTA 398
DB 250 QFAPDTFLNVLILFLETNPANLINGIGLAVGLFLAGQIKSLFPMNTYFIAMTRVGAKIQTM 309
QY 399 VIGAVYKALVITNSAKSSVGEIVNLMSVDAQRFMDLATYINMTSAPLOVIALYLL 458
DB 310 LSCAVYKESLLSTNARERTVGEVMNLTSLIDVRFMTIQIQQYMSSPQIILICVYL 369
QY 459 WLNIGPSVLGAVAVWVLMVNVNAVMAMKTQYOVYAHMSKDNRIKLMEIINGIVLTKLY 518
DB 370 SQTVGAVAWAGIIVMISTVPIINICSMITTKWQILMKYKDBRRLINEVINGIKIVKLIS 429
QY 519 AMELAFKDKVLAIQEELKVLKKSAYLSAVGTFTWCTPPLVALCTFAVYVYTIIDENIILD 578
DB 430 AMETRMETTERVADKELKMKIKQSALKTFPADCLNVGAPVYVALLSFVFLPIDKXVLT 489
QY 579 AQTAFVSLAFNLIIRFPLNIIIPWVSIISIVQASVSLKRLIFLSHEELEPDSIERRPVKG 638
DB 490 PNIFVSLSTFNLRGFLMMAAEFLVACVQLVNSKRVRTFLCEKEVDTAIDKE--IRGE 548
QY 639 GGTNSIIVRNATFTWASADPPTLNGITFSIPEGALVAVVGOVGGCKSSLNALLAEMKY 698
DB 549 LYTIVTEHSSGSPAMDSEARILSDIEFLASKELLYTVGVSGVSKSSLNALLAAGEMKY 608
QY 699 EGHVAIKSVYVYVPOQAWIIONDSLRENTLPFCOLEEPPYRSYVIOACALLPDEILPESGR 758
DB 609 CGYGVGRSVAYILSQPILNOSLKNVLMQADLNDVLYKVIYISSCALKEDEKQLPDGD 668
QY 759 TEIEKGYNISGGQKORVSLARAVYSNADIVLPDDPLSAVDAAHYKHIIFENVIGEKMLK 818
DB 669 TEIEKGYNISGGQKARIALARAVYQSKDYVFLDPLSAVDAAHYKHIIFENVIGENKML 728
QY 819 NKTILVTHMSYTPQVDVIVYMSGGKISENGSTQELLARQAPAEPLRTAYASTQEOEDA 878
DB 729 HTTILVLTNCSPFLQESGKIIVMKG-----ETRYESSG 761
QY 879 EENGVATGSGPGKAKOMENGMLVTDASGKQLOROLSSSSSYSGDISRHHNSTLEQAE 938
DB 762 EESG-----GEENSDFLPQSTIASGSMKSLR--LSK-----ISKKSKSSTVEK-- 804
QY 939 AKKEETWKLMEADAKQTGOVYLIVYWDYKAIQGLFISFLIFLPMCHVS--ALASNYLS 997
DB 805 -KKQDA--LTKKEAALIGRVNPGVYLLYFKAMGIIVTYVLPALIAVLNVSPALGRSLWLT 861
QY 998 LMTDDPLVINGIQEHTKY--RLSYTGALGISGCIIVFGKSAVNSVIGGILASCHAVDLHS 1055
DB 862 AMSANDIDINHPDTMSVGARLVAGFQITFVILFSLVLLLLIGVAAASNTLKHKPLHAN 921

1088 DSISRSFVYAQFGEALNGLSTIRAYKAYDRMADINGRSMNNINFTLVNMGANFGLRL 1147
QY 1204 ECVGNICIVLFAALFAVI-----SRHSLSAGLVGHSVSYSLQVTTYLMVLVMSSEMETNI 1258
D 1148 ETLGGLMILWLTASAVVONGAENQOAFSTMGLLSTALNITSLTGLSVLRASLAEHSL 1267
QY 1259 VAVRLKEYSETKEAPMOIOETLAPSSWPQVGRVFRNVCRLRYREDIDFVLRHINVTIN 1318
D 1208 NAVRVRGVYIEIPREAPRVINENRPPGMPSSGSIKEFDVVLARYPOLPRLVHVSFEFH 1267
QY 1319 GGEKVGIVGRTGAKSSLTGLFRIINESAEGEIIIDGINAKIGHDIFRITIIPODPV 1336
D 1268 PTDKVGIVGRTGAKSSLTGLFRIIVEYKGRILIDCDGKFGGLMDIRKVLGIIPSPV 1327
QY 1379 LFGSGLRMMNDLPFGQSDSEEVWTSLELAHLKDFVSALPDKIDHECAEGENLSYGORQV 1438
D 1328 LFGSTVAFNIDPFGHNDADIMESIERHAKDTIRRNPLGIDABVSEGENFSGOQRL 1367
QY 1439 CLAPALIRKYLIVLDETAAVLDETDDLIOSTIRTOGEDCTVLTIAHRLNTMDYTRVI 1498
D 1388 SLSPALIRRSKILVLDETAADVTDALIOKTIREEFKSCMTMLIAHRLNTIIDCKIL 1447
QY 1499 VLDKGEIOEYVAPSDILQORGLFYS-MAKDAG 1529
D 1448 VLDSGRVOEFSSPENLISNEGSSPFSKWOVSTG 1479

RESULT 12

D86428
glutathione S-conjugate transporting ATPase (AtMRP1) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: D86428
R/Theologian: A.; Ecker, J.R.; Palm, C.J.; Federgruel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Matli, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: D86428
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1622 <STO>
A/Cross-references: UNIPROT:Q9C8G9, UNIPARC:UPI000009CF47, GB:AE005172, NID:g11055814; F
C/Genetics:
A/Map position: 1
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 23.8%; Score 2356.5; DB 2; Length 1622;

Best local similarity 35.6%; Pred. No. 2.9e-151;
Matches 569; Conservative 274; Mismatches 564; Indels 193; Gaps 31;

QY 47 MACPPF-YFLKSHDRGIOMTPIANKTKLGL-----FLMTVCMADLFYFW 94
D 8 WYCPVPNGVMTKTVDVAFGAYTPCAIDSPVGLISHLVLLILCLVRLMLITKHKVDFC 67
QY 95 ERSR-GIFLA-----PVFLVSPTLIGITTL-----LATFLIQLEKRGQS 134
D 68 LRSMFSPFLLLAAYATAEPLFRU---VMRISVLDLDGAGPRPEALMVLVE-AFAMG 122
QY 135 SGIMLTF-----WLVALVCALALIRSKIMTALKEDAQVDLDFDITFYVYFSL 183
D 123 SALVMTVETKYTHIELRMVYRFAVIALVGMVLMNLVLGVK-EYGSFKLXIYISVA 181
QY 184 IQLV-----LSCFSDRSLPFSFTIHD-----PNPPESSASGLSITITWW 223
D 182 VQVAFGLTLLFVYFPNLDIPYGVTTVPVGTENSEDEYEEELPGENICPERHANLPDSIFPSW 241

QY 224 ITGLIYVGRPOPLGSDPLMSLNKEPDSQVPPVLVKNWKECAKTRQPVKVYSSDPA 283
D 242 LNPMLTIGSRKPLTERKXVWHLDTWKDETETLRSPQKSMDELEK----- 285
QY 284 QPKRSSKVDANEEVEALIVSPQENRPSLEKLYKTFEGYPLMSFFPKAIHDMFSGP 343
D 286 -PK-----PMLRALNLSLGGFRMGWGKIKGNDSCQVGP 320
QY 344 QILKLIRFVNDTKAPDMQGYFYTVLLFVTAQTLVLAHQYFHLCFVSGMRKIKTAVIGAV 403
D 321 LLNLELLKSM-QNEPAMIGIYALISIFGVVIGVLEAQYFQVWVGVRLRSALIAAV 379
QY 404 YRKALVTINSARKSSTGYEIVNLMSVDQRPMDLATTYNNIWSAPLOVILALYMLNLG 463
D 380 FRKSLRLTNGRKKFQFGKITMLMTTDAESIQQICQSLHTMWSAPFRILVALVLLYYOOLG 439
QY 464 PSVLGAVVAWVAWVAWVNAAMKTKTYQVAMHSKNDRIKMLNLIINGIKVLKLYAMELA 523
D 440 VASTIGALFLVLMPIOTVITISKTQKLTKEGLOTKRIGLMEVLAAMDITKCYAMENS 499
QY 524 FKDKVLAIROBELKVLKKSAYLSAVGTFTWCTPFLVACTFAVYVTIDENNIIDAQTA 583
D 500 FQSKVQVTRDDELSPFRKAQLLSAFNMFLNSIFVLTVVSVGVSLGGD-LTPARAF 557
QY 584 VSLALFNILRPPLILDMVTSISVQASVSLKRLRIFLSHEE--LEPDSIERRPYKQGGT 641
D 558 TSLSPFSLRFPFLMFLNIIITQWVANVSINRLEEVLSIERVLLPNP---PIEP--CQ 611
QY 642 NSITVRNATFTW-ARSDPPLTNGITFSIPGALVAVVGVGCGKSSLSLALAMD-KVE 699
D 612 PAISIRNGVSWDSKARPTLSNINLDIPQSLAAVVGSTGEGTSLISALMGLPARSD 671
QY 700 GHVAKISVAVVPOQAMIONDSLRENTLFCQGLEEPPYRSYIQAICALPLLEIPSGDR 759
D 672 ATVLRKSAVAVPVQSVIIPATVRNDLILFAPRQEKERYVIDTALQHDLELLPGDILT 731
QY 760 EIGKGVNLGGQKQVSLARAVYSNADIVLFDPLSAVDAVHGKHIENVYIGKGMKN 819
D 732 EIGRGVNIISGQKQVSMARAVYSNDVICLIDPLSALDAHQVQVEKCI--KREIGQ 789
QY 820 KTRILVTHMSKYLPOVDVIVIMSGKSIEMGSIYELLARQCAFAPLRTASTQEQDAE 879
D 790 TTRVLVNLQHLFSLQVKKILVHEGYKEBGTVELCHSGPLFRLMENAKVDY--SE 847
QY 880 ENGVTVSGPGEKAKOMENGLVTDASAKQLQORQLSSSSSYSGDISRHNSHSTAELQ-- 936
D 848 ENGEAEVD--QTSKVPENG-----NANLQXDGI 875
QY 937 -AEAKKEETWKLMEADRAQGTQGVKLQVYWDYMKAI-GLFISFLSIFLMCNHVSALSNY 994
D 876 ETKSKSGNSVLYRREERTGVSMKVLERYQNALGAMVVMVLVICVYLQVFRVSSST 935
QY 995 WLSLMTDD--PIVNGTQEHKTVRLSYGALGISCIAVFGSMAVSIIGILASCIVLD 1052
D 936 WLSMTGTSGPTKTHGPFPYN---IVYALLSPGVSVTLINSYWLIMSSLYAAKMDAM 991
QY 1053 LHSILRSPMFEERTPGCNLVNRFSEKLDYDWSMIPYIKFMGSLFNVIGACIVILIT 1112
D 992 LGSILRAPMWFQPNPLGRILINRPAKOMGIDIRVAVFANMFQSIQQLSTVLIGIVS 1051
QY 1113 PIAAIIIPGLILYFVQRPVVASRQIKRLSEVSRSPPVSHFNETLIGSVIRAFEOE 1172
D 1052 TSLMALIMPILLVGYAVLYQNTSRREIKMDSITRSVVAQAQFDEALNGASSIRAYXD 1111
QY 1173 RFIHQSDIKVDENQKAYYPSIVANRMLAVRLECVCNCIVLFAALFAVI-----SRHSLSA 1227
D 1112 RMAEINGRSMNNIRFTLVNNAANRMGLIRLEVLVGGMLVMTLTAJAWQNGKAAQOAYA 1171
QY 1228 GLVGLSVSYSLQVTTYLMVLVMSSEMETNI-VAVERLKEYSETKEAPMOIOETAPPSW 1267
D 1172 STMGILSLYSTLSTSSLTSLVRLASLAEHSLNSVERGANTYIEISEAFLVLENRPPGWM 1231
QY 1288 PQVGRVFRNVCRLRYREDIDFVLRHINVTINGEKVGIVGRTGAKSSLTGLFRINESA 1347

Db 1232 PSSGSIKFEEDVLAAREDELPEVLHGVFLISPMOKVGI VGRGTAGKSKSLNALFRIYELE 1291
Qy 1348 EGEIITIDGINIAKIGLHDLAFKTIITIPDPLFSGSIRMLNDPFSQYSDERWTSLELAH 1407
Db 1292 KGRLLIDECIGIRGLMDLKRVLGIIIPQAPVLFSGYRFPNIDPFSRHNDLWESLRAH 1351
Qy 1408 LKDFVSAIPDLDECAEGENISVGGOROLVCLARALLRKTLVLEATPAVDLETDDL 1467
Db 1352 LKOTIRNPVGLDPEVTEAGENSVGOROLISLARALLRSKILVLEATPAVDRTDL 1411
Qy 1468 IOSTIRTOFEDCTVLTIAHRLNTIMDYTRVILDKGEIOEGYASPDLLQORGLFYSNAKD 1527
Db 1412 IQOTIREFESCTMLIIAHLNLTIIDCDKYLVDLSKGVGFESSPENLISNGESSFSK--- 1468
Qy 1528 AGLVGGGGGMLSRKGIIPBEYVITRLAEDPAERFYTRE 1567
Db 1469 -----MVOSTGTANAAYLRSLITLENK-----RTRE 1493
RESULT 13
E86428
Probable ABC transporter [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: E86428
R/Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Nansen, N.F.; Hughes, B.; Hlizar, L.
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiz, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: E86428
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-1495 <STO>
A/Cross-references: UNIPROT:Q9C8H0; UNIPARC:UPI0000482FF; GB:AE005172; NID:G1055818; F
C/Genetics:
A/Map position: 1
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
Query Match 23.5%; Score 2326; DB 2; Length 1495;
Best Local Similarity 34.3%; Pred. No. 3e-149;
Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;
Qy 2 ALNFGCS--ADGSDPLMDANVTNTSNDFTCFQNTVLVWP-----CFY-LWACFP 51
Db 5 ALNWCYCPVADG-----WEKAVDGAFAVTPCAIDSLVMLVSHFVLGLCFRIWITIF- 58
Qy 52 FYFLVLSRHRGRTQMTPLNKTALGFLMLVYCNAD-----LFSFMRSGIFLAPV 105
Db 59 -----HNTKAOIYVLRKKYNCVIG-LIACYCVAEPVRLVWGSLFPMDEETDPFP 110
Qy 106 FLVSPITLIGTTLATFLQLERKRGVSSGIMLTFVLVLCALAIIRSKINTALKEA 165
Db 111 EVASLMEVFAWMSMLVIGLETKQYKFRWYRVRGVLYVADAVLDLVLPKRSIN 170
Qy 166 QVDFRDT---FYVYFSLILQLVLSGFSRSPPLFSETITHP-----NPCP 209
Db 171 RTLVYFISGRCSQALFGILLIYIPEL--DYPGYHIVNNEPLDNEVDALGGEHICP 228
Qy 210 ESSASFSLRTFMWITGLYRGYROPLESGSLWLNKEDTSQGVNPLVNMKKECAKTR 269
Db 229 ERHASFISRIYFMWITGLYRGYROPLESGSLWLNKEDTSQGVNPLVNMKKECAKTR 286
Qy 270 KQPVKVVYSSKDDPAQFESSKVDANEVEALIVYSPQKWNPSLFLKVLTYTFGPFYLMSP 329
Db 287 -----PK-----PVLRLALNNSLGGRTFLTAG 307

Qy 330 FFAKIDLMWFGSPQILIKLFVNDTKAPDMQGYFYVLLFYTACQLTVLHQYHICF 389
Db 308 IFKIGNDLSQFVGPVILSHLRSMQBGD--PAMGYVAFIIFPGVTGLVCEBQYFQNVW 366
Qy 390 VSGMRKTAIVAGVYRKALVITNSARKSSTVGEIVMLMSYDAORFMDLATYINMWSAPL 449
Db 367 RVGFRLRSTLVAALFHKSLRLTHEARKNFASGAVTMTITDANAALQOISQHLGWSAPF 426
Qy 450 QVILAYLLMLNIGPSVLGAVAVMIVPVNAVAMKTKTYQVAHMSKDNRIKLMNEIL 509
Db 427 RIIVSMILLYQQGVASLFSLSLILFLIPLQTLIIISMRKLTKEGLQMTDKRVTWEL 466
Qy 510 NGIKVLKLVWELAFPDKVLATIQEELKYLKSAVYLSAVTFPWCCTPFLVACTPAPVY 569
Db 487 SSMDTVKCYAWESFESRIQGINNEELSWERKQALLAFNSFLINSTPVAVTVASFEVVF 546
Qy 570 TIDENNILDAQTAFAVSLALFNILRPPLNIPWYISSIVQASVSLKRL-RIFLSHELEDP 628
Db 547 LIGGD--LTPARAFSLSPFAVIRFPPLNIPILISOVNAVAVSLQRIEELLSEERILAQ 604
Qy 629 SIERRPVKGGGNTSITVRNATFTW-ARSDPTLNGITFSIPGALVAVVGVQVCGKSL 687
Db 605 NPLQF-----GTPALSIKRGVSMDSKTKPTLSDINIEIPVGTILVALVGTGEGKSTL 659
Qy 688 LSLALLBMDKVE-GHVAIKGSVAVYVQOAMTQDSIRENLEFRCQLEBPYRYIOACAL 746
Db 660 ISMLDELSHAETTSVIRSVIRSVYVPOVSVIFNATVRENTLFGSDDESERYMRADITATL 719
Qy 747 LPLELILPSPDRTEIGEKGNLSGGOKORVSLARAYSNADYLPDPPLSAVNAVHKKH 806
Db 720 QHDLDLPGDLTEIGRGNVSGGOKORSMARAYSNADYLPDPPLSAVNAVHKKH 779
Qy 807 FENVIPKGMKKNKTRILVTHSMYLPDYDVILVMSGKISENGSOYELIARDGAPAEFL 866
Db 780 FDSGM--KDELKGTHTVLTVNTQHLPLMDKILVSEGMIKEGTVEILSKSGLIFKGM 837
Qy 867 RITYASTQEBQDAENGVTGSGPGKAKQENOMLVTDSAGKOLQOLSSSSSGDISR 926
Db 838 ENAKMDATQEVNTN-----DENLIKGPPTVTVVS--ERNLGSTK----- 876
Qy 927 HNSTALELQAEKKEETWMEADRAQGVQLSYWMDPKAI-GLTFSLSFILPMCN 985
Db 877 -----QKRRRS-VLTKQEBREGLIISWNLWKYKAVGGLVMMILACTIAT 924
Qy 986 HVSALASNYWLSLMTDPIVNGTQEHK-----VRLSVYGALGISOGIAVFGYMAVSIG 1040
Db 925 EYLAVSSSTWLSITWD-----QSTSKVYSPGYIVVVALGFGQVAVFTNMFMLITS 977
Qy 1041 GILASCLHVDLHLSLRSPMSFFERTPGNVLNRSKELDTVDSMIPEYIKMFGSLEN 1100
Db 978 SLHAARRLHDMSSILRAPMLFFHTNPTGRVYNRESKDIGIDRVANVAMNNFNAQIMQ 1037
Qy 1101 VIGACTVILLATPIAIIIPRLGIYFPVQRFVYASSROLKRLSRSRGSYHFMETLL 1160
Db 1038 LSTFALIGTVSTISLWAMPPLILFYAALVLYQSTSRERRLDSVTRSPIYAQFEBALN 1097
Qy 1161 GVSIVAFEEQERFIHQSLDKVDENOKAYYSIVARWMLAVRLGCGNCIVFAALFAYI 1220
Db 1098 GLSSTAYKAYADMAINKSKMDNNTIRPLANTASSRMLTIRLETIGGWMITLAFAYL 1157
Qy 1221 SRHSLS-----AGVLGSVYSYLOVTVYVNLVMSSEMETNIVAVERLKEYSETEKAP 1275
Db 1158 QNENTNNOAGFASGTWGLISLTINTISLSGLVLRQASRAENSINSVERVANYIDLPESEAT 1217
Qy 1276 WOIOETAPSSPQVQVREFRNYCLAREDDLPVLHINVTINGCKGVIGRTGAGKS 1335
Db 1218 DIIENNRPVCGMPDGGSTIRFEDVHLRPGLPVHLGLTFVYSPSEKVGAVGTGAGKS 1277
Qy 1336 LTLGLFRINESAGEIITIDGINIAKIGLHDLRFKTIITIPDPLFSGSIRMLNDPFSQY 1395
Db 1278 MNALFRIYVEVEGRIMDDCVAKRGLDVARVSIIFQSPVLFSGYRFPNIDPFSH 1337

A:Residues: 1-255 <BOY>
A:Cross-references: UNIPARC:UPI00001746CF
A:Note: the authors translated the codon CAG for residue 248 as His
R:Purnelle, B.; Goffeau, A.
Submitted to the EMBL Data Library, April 1996
A:Description: The sequence of 32 kb on the left arm of yeast chromosome XII reveals 14
mily and a new ABC transporter homologous to the human multidrug resistance protein.
A:Reference number: S69380
A:Accession: S69391
A:Molecule type: DNA
A:Residues: 1-1559 <PDR>
A:Cross-references: UNIPARC:UPI0000126A9C; EMBL:X91560; NID:g1297003; PIDN:CAA62162.1; F
R:Mosga, T.; Zimmermann, F. K.
Yeast 12, 693-708, 1996
A:Title: Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on a 43.7 kb
e conductance regulator protein CTR.
A:Reference number: S70557; MUID:96405918; PMID:8810043
A:Accession: S70560
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1559 <MW>
A:Cross-references: UNIPARC:UPI0000126A9C; EMBL:X91488; NID:g1495203; PIDN:CAA62776.1; F
C:Genetics:
A:Gene: SGD:BPT1; MIPS:YLL015W
A:Cross-references: SGD:S0003938
A:Map position: 12L
A:Superfamily: human multidrug resistance protein CMOA2; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F:28-44/Domain: transmembrane #status predicted <TM1>
F:143-159/Domain: transmembrane #status predicted <TM2>
F:178-194/Domain: transmembrane #status predicted <TM3>
F:334-350/Domain: transmembrane #status predicted <TM4>
F:421-437/Domain: transmembrane #status predicted <TM5>
F:526-542/Domain: transmembrane #status predicted <TM6>
F:550-566/Domain: transmembrane #status predicted <TM7>
F:654-679/Domain: ATP-binding cassette homology <ABC1>
F:672-679/Region: nucleotide-binding motif A (P-loop)
F:974-990/Domain: transmembrane #status predicted <TM8>
F:1017-1033/Domain: transmembrane #status predicted <TM9>
F:1099-1115/Domain: transmembrane #status predicted <TM10>
F:1118-1134/Domain: transmembrane #status predicted <TM11>
F:1212-1228/Domain: transmembrane #status predicted <TM12>
F:1319-1529/Domain: ATP-binding cassette homology <ABC2>
F:1336-1343/Region: nucleotide-binding motif A (P-loop)

Query Match 23.2%; Score 2294.5; DB 1; Length 1559;
Best Local Similarity 35.1%; Pred. No. 4.4e-147;
Matches 574; Conservative 292; Mismatches 541; Indels 227; Gaps 39;

QY 23 NTSNPDTKCFONTLVVWPCFYIM-ACFPYFLYSRHDGYIQMTPLNKTALGFL- 80
DB 23 NALNP-----CFISVISAMQAVFLLISGYOLMKLYKNNKVPKPKNPTLPKSKINSHLT 78
QY 81 -LWIVCW-----ADLFYSFMRSGIFLAPVPLVPLLTG---TTLAT 121
DB 79 HLTNVCQSTIIICELALVSOSSDRYPFLIK-KALYLNLF-----NLGSLPTQYLA 132
QY 122 FLIOLERRKVOSSGIMLTFWLVALVCAALIRSKI-----MTALKEDAOV 167
DB 133 F-----KSTFSNGNOLFYYMPOILLQFLILORRYHSSNERTLYISQGTAMILEVL 185
QY 168 DLPRDITFYVYFSLILQVLVSCFSDRSPLFSETIHDNPPCESSASFLSRITFWITGL 227
DB 186 -LFPNSVAFIY-DLCIFEPI-----NELSEYKKNGWY-----PVHVASITFTIMNKL 233
QY 228 IYRGYRPLRGSDLMSTLTK--EDTSEQVAVLVNQMKECAKTRKQVKKVYSSKDAOP 285
DB 234 IVEIYR-----NKKIKDPNQPLPVPDLNLIK----- 259
QY 286 KESSKVDANEVEALIVKSPQKEWNPFLFKVLYTFTPPYFLMSPFKAIHDLMMFSGPOI 345
DB 260 SISKEFKANWELKMLNRN-----SLMRAIWKSPRTISVAMLVETTSDLISLVQPOF 312

QY 346 LKLIKRFN---DTKADMGQFYTVLLFTYACQLTVLHQYHICFVSGMRITNAVIG 402
DB 313 LRIPIDGLNPBTSSKYPPLNGVFTALTFLVISVSVFTNQFYIGITEAGIGIGIASL 372
QY 403 VYRKALVITSARSKSTVGEIVNMSVDA---GRFMDLATYNNIMWAPLOVIALYLLW 459
DB 373 VYQSLTLTLAERNEKSTGDLINMSDVLRIGQFFENA---QITIGAPIOIIYVLSLV 429
QY 460 LNLGSPVLAVVAVVAVPVNAVAMKTKYQVAHMSKONRIKLMNBIINGIKVLKYA 519
DB 430 WLKGAVIGGLVWVAIMVPIAFSLRKYKLSKTOMKYKMRIRKITELLNAIKSIKYA 489
QY 520 WELAFKQVLAIRE-ELKYLKSAVLSAVGTFTWCTPPVLVACTAAYVTTIDBNILD 578
DB 490 WEEPMARLHVNRDMLKNFRKIGVSNLIYFAMNCVPIPMVCTSGLP-SLPSDPLS 548
QY 579 AQAFVSLATFNILRFPNLILPMVISIVQASVSLKRLIFLSHEELPDSIER-RPVKD 637
DB 549 PALVPSLSPLNLSAIVSVPSMINTIITSVMERLKSFLSDEIDDSFTEIRIDSD 608
QY 638 GCGTNSITVNAFTW-----ARSDPTLNGIT-----FSIPEGALV 674
DB 609 ERALPALEMNNITFLMKSKEVLTSSQGDMLRTDESEIIGSSQALKNIDHFEAKRGD 668
QY 675 AAYGVQVCGKSSLSALLAEMDKVEGH-----VAIKSVAYVPOQAMIQNDSLRENI 726
DB 669 CVGRVAGKSTFLKALIGQLPCMSGSRDIPPLKLIIRSSVAACOSBSMTMNASVRENI 728
QY 727 LFGQLEEPYRVSIVQCALLPLEILPSDPRTEIGEGVNLGGQKQVSLARAVVNSA 786
DB 729 LFGHKPDQDYDLTKAQQLPDKLIPDGBETLVSGKGLSSGQARSLAARAVSRA 788
QY 787 DIYLFDDPLSAVDANVGHKIFENV-IGPKGMKNKTRILVYTHSSVYLPQVDVITVMSGGK 845
DB 789 DYLDDIILSAVDAEVSKNIEEVLICKTALNNKTIILTNVYSLIKSGMIALNGE 848
QY 846 ISEKSGYQELIAR--DGAFAELRTYAS---TEQDQAEANGTVGSGPKAEAYOME 897
DB 849 IVEQGNEDVWNRKNTSKLKLLEEDPSIDNGESDVQTEHRSSEVBP----- 900
QY 898 NGMLVDSACKOLQROUSSSSVSGDISRHNHNSPAELQXBAKKEFWKL----- 947
DB 901 -----IQLKATSEETDEVT---ESELRIKNSRASIATLRPRPFGAQL 945
QY 948 -----MEADKAQTQGVKLVSVMYDWYKATGLFISFLIFLM-CHNVSALASNYSLW 999
DB 946 DSVKTAQKAKTEKEVGRVKTIYAIYKAGV-LGVLFPLMLTRVPDLAENFWLKYW 1004
QY 1000 TDDPIYNGOEHTRKVLSTYGALGISQGIAPFGYSNAVSI--GGIILASCLHVDLHST 1056
DB 1005 SESENEKSGNERWMMFVGVSILGVAS--AAFNNLASIMMLVCSIGSKKLEHSMKSV 1062
QY 1057 LRSPMSFEETPGNVLNFRPSKELDPTVDSMIPYIKMFGSLFNVGACIVILATPIA 1116
DB 1063 IRSPMTFEETPGRLNFRSSDMDAVDSMLQYIPSEFFKSLTYLTVTLVGVNMPWL 1122
QY 1117 IILPGLILYFYQRFVAVSARQLKRLSRSRSPVYSHFNETLIGVSVIRAFEOERFIH 1176
DB 1123 VFMFLVVIITYQTPIYVLSRELKLIISTYSPIMSLMSESLNGISIIDAYHFERFIY 1182
QY 1177 QSLDKVDENOKAVYPSIVANRWLAVERECVGNCTVLPALFAVI---SRHSLSAGVLGS 1233
DB 1183 LNYEKTQYNDVFENFESTRMRSLVRLQITIGATVATATATATATAMTKQLSSGMVGL 1242
QY 1234 VYSLOYTTLNMLVRRSSSEMETNIVAVELEKYSSEBEKAPMOIQETAPPSPOVGRV 1293
DB 1243 MSYSLEVTGSLTIVRTTIVETINIVSVERIYVCELPAPQASINBEKRDENKPSKGI 1302
QY 1294 EFRNVCYLRVEDLDVFLRHINVTINGEKVYIGRTGAKSSLTGLFRINESAGEFIIT 1353
DB 1303 EFKVSTGYKRENDPVLNNINVKIEPCEKGVIGRTGAKGSTLALFRILEPTEGKII 1362

QY 1354 DGINIAKIGLHDLRPFKTTIIPQDPVLFSGSLRMLNLDPSOYSDEBWTSLFLAHKDFVS 1413
 Db 1363 DGIIDISDGLFDLRSHLAIIPQDAQAFEGTVKTNLDPFNRYSEDELKRAVEQAHKPHLE 1422
 QY 1414 AL-----PDKLDHECAEGENLSVGOROLVCLARALLRKTILVLDIA 1456
 Db 1423 KMLHSGPRGDDSNEDGNVNDILDVKINENGSNLSVGOROLLCLARALLNRSKILVLDIA 1482
 QY 1457 TAAVDLETDLDLIGSTIRTOFEDCTVLFTHRLNTIMDYTRYVLDPKGEIOEYGAPSDLLQ 1516
 Db 1483 TASVDMETDKIIGDTIRREFKDRITLITIAHRIDTVLDSDKIIVLDQGSVREFDPSKLLS 1542
 QY 1517 OR-GLFYSMADAG 1529
 Db 1543 DKTSIFYSLCEKCG 1556

Search completed: December 15, 2005, 15:36:42
 Job time : 54.8285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:56:46 ; Search time 243.004 Seconds
(without alignments)
5594.767 Million cell updates/sec

Title: US-10-665-283-6

Perfect score: 9901

Sequence: 1 MALRGFCSDAGSDPLMDMNV.....ASVAVAKAKPKFSISPSLS 1927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_05.80.*
2: uniprot_sprot.*
3: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7860	79.4	1531	1	MRP1_HUMAN
2	7769	78.5	1515	2	MRP1_HUMAN
3	7695	77.7	1531	2	MRP1_HUMAN
4	7689	77.7	1531	2	MRP1_HUMAN
5	7444	75.2	1459	2	MRP1_HUMAN
6	7419.5	74.9	1456	2	MRP1_HUMAN
7	7272	73.4	1531	2	MRP1_HUMAN
8	7242	73.1	1439	2	MRP1_HUMAN
9	7230.5	73.0	1530	2	MRP1_HUMAN
10	7104.5	71.8	1400	2	MRP1_HUMAN
11	7002.5	70.7	1528	1	MRP1_MOUSE
12	6932.5	70.0	1532	2	MRP1_MOUSE
13	6932.5	70.0	1532	2	MRP1_MOUSE
14	6892	69.6	1523	2	MRP1_MOUSE
15	6801	63.6	1525	2	MRP1_MOUSE
16	6075	61.4	1215	2	MRP1_MOUSE
17	4526	45.7	1522	2	MRP1_MOUSE
18	4522	45.7	1519	2	MRP1_MOUSE
19	4522	45.7	1522	2	MRP1_MOUSE
20	4508.5	45.5	1523	2	MRP1_MOUSE
21	4491.5	45.3	1527	1	MRP1_HUMAN
22	4484.5	45.3	1533	2	MRP1_HUMAN
23	4422	44.7	1514	2	MRP1_HUMAN
24	4402	44.5	1498	2	MRP1_HUMAN
25	4396	44.4	1522	1	MRP1_MOUSE
26	4148.5	41.9	1496	2	MRP1_MOUSE
27	3885.5	39.2	1548	2	MRP1_MOUSE
28	3880.5	39.2	1548	2	MRP1_MOUSE
29	3879.5	39.1	1548	2	MRP1_MOUSE
30	3874.5	39.1	1548	2	MRP1_MOUSE
31	3860.5	39.0	1548	2	MRP1_MOUSE

32	3857	39.0	1549	2	MRP1_MOUSE	07ktc0_drosophila
33	3845.5	38.8	1548	2	MRP1_MOUSE	08t9c5_drosophila
34	3838.5	38.8	1548	2	MRP1_MOUSE	07ktb8_drosophila
35	3836.5	38.7	1548	2	MRP1_MOUSE	07ktb9_drosophila
36	3818.5	38.6	1548	2	MRP1_MOUSE	07ktc0_drosophila
37	3813.5	38.5	1548	2	MRP1_MOUSE	07ktc8_drosophila
38	3812.5	38.5	1548	2	MRP1_MOUSE	07ktc4_drosophila
39	3807.5	38.5	1548	2	MRP1_MOUSE	07ktc9_drosophila
40	3793.5	38.3	1548	2	MRP1_MOUSE	07ktc7_drosophila
41	3790	38.3	1549	2	MRP1_MOUSE	09vks6_drosophila
42	3771.5	38.1	1548	2	MRP1_MOUSE	07ktc5_drosophila
43	3769.5	38.1	1548	2	MRP1_MOUSE	07ktc6_drosophila
44	3741	37.8	1564	2	MRP1_MOUSE	08q9c8_raja_erinac
45	3738.5	37.8	1567	2	MRP1_MOUSE	06ph26_brachydanio

ALIGNMENTS

RESULT 1
MRP1_HUMAN STANDARD; PRT; 1531 AA.
ID MRP1_HUMAN
AC P3527; O14819; P78419;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).
GN Name=ABCC1; Synonyms=MRP, MRP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE. PubMed=1360704;
RX MEDLINE=93088080; PubMed=1360704;
RA Cole S.P.C., Bhargava G., Gerlach J.H., Mackie J.E., Grant C.E., Almqvist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;
RT "Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line."
RL Science 258:1650-1654(1992).
[2]
RP SEQUENCE REVISION.
RX MEDLINE=93262415; PubMed=8098549;
RA Cole S.P.C., Deeley R.G.;
RT "Multidrug resistance-associated protein: sequence correction."
RL Science 260:879-879(1993).
[3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96251691; PubMed=8649356;
RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;
RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter."
RL Mol. Pharmacol. 49:962-971(1996).
[4]
RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA."
RL Genomics 45:368-378(1997).
[5]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA) OF 1131-1531.
RX MEDLINE=99425270; PubMed=10493829; DOI=10.1006/geno.1999.5927;
RA Loftus B.J., Kim U.-J., Shedd V.P., Kalush P., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.B., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q."
RL Genomics 60:295-308(1999).

[6] TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=97442425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;
 RA Higginer D.R., Almqvist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,
 RA Deeley R.G., Cole S.P.C.;
 RT "Membrane topology of the multidrug resistance protein (MRP). A study
 of glycosylation-site mutants reveals an extracytosolic NH2
 terminus.";
 RL J. Biol. Chem. 272:23623-23630 (1997).
 RN [7]
 RP TOPOLOGY.
 RX MEDLINE=97476249; PubMed=9334225; DOI=10.1074/jbc.272.42.26479;
 RA Kaet C., Gros P.;
 RT "Topology mapping of the amino-terminal half of multidrug resistance-
 associated protein by epitope insertion and immunofluorescence.";
 RL J. Biol. Chem. 272:26479-26487 (1997).
 RN [8]
 RP TOPOLOGY.
 RX MEDLINE=98153110; PubMed=9485377; DOI=10.1021/bi972332v;
 RA Kaet C., Gros P.;
 RT "Epitope insertion favors a six transmembrane domain model for the
 carboxy-terminal portion of the multidrug resistance-associated
 protein.";
 RL Biochemistry 37:2305-2313 (1998).
 RN [9]
 RP MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-GLU-1455.
 RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;
 RA Cui L., Hou Y.-X., RJordan J.R., Chang X.-B.;
 RT "Mutations of the Walker B motif in the first nucleotide binding
 domain of multidrug resistance protein MRP1 prevent conformational
 maturation.";
 RL Arch. Biochem. Biophys. 392:153-161 (2001).
 RN [10]
 RP MUTAGENESIS OF TRP-1246.
 RX MEDLINE=21238219; PubMed=11278867; DOI=10.1074/jbc.M011246200;
 RA Ito K., Olsen S.L., Qiu W., Deeley R.G., Cole S.P.C.;
 RT "Mutation of a single conserved tryptophan in multidrug resistance
 protein 1 (MRP1/ABCC1) results in loss of drug resistance and
 selective loss of organic anion transport.";
 RL J. Biol. Chem. 276:15616-15624 (2001).
 RN [11]
 RP VARIANTS GLN-633 AND VAL-671.
 RX MEDLINE=20296630; PubMed=10835642; DOI=10.1038/76102;
 RA Le Saux O., Urban Z., Tschuch C., Csiszar K., Bacchelli B.,
 RA Quaglini D., Paeguali-Ronchetti I., Pope F.M., Richards A., Terry S.,
 RA Bercovich L., de Paeppe A., Boyd C.D.;
 RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma
 elasticum.";
 RL Nat. Genet. 25:223-227 (2000).
 RN [12]
 RP VARIANT VAL-671.
 RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;
 RA Ringpfeil F., Lebowitz M.G., Christiano A.M., Uitto J.;
 RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a
 transmembrane ATP-binding cassette (ABC) transporter.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006 (2000).
 RN [13]
 RP VARIANT SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.
 RX MEDLINE=21578494; PubMed=11721865; DOI=10.1007/s100380170017;
 RA Corrad S., Kuffmann H.-M., Ito K., Deeley R.G., Cole S.P.C.,
 RA Schrenk D.;
 RT "Identification of human multidrug resistance protein 1 (MRP1)
 mutations and characterization of a G671V substitution.";
 RL J. Hum. Genet. 46:656-663 (2001).
 RN [14]
 RP VARIANTS MET-117 AND LEU-1512.
 RX MEDLINE=20579883; PubMed=11139250;
 DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14>3.0.CO;2-F;
 RA Perdu J., Germain D.P.;
 RT "Identification of novel polymorphisms in the pms and MRP1 (ABCC1)
 genes at locus 16p13.1 and exclusion of both genes as responsible for
 pseudoxanthoma elasticum.";
 RL Hum. Mutat. 17:74-75 (2001).

[15]
 RP VARIANTS SER-433; ILE-73; GLN-723 AND GLN-1058.
 RX MEDLINE=21163848; PubMed=1126082;
 DOI=10.1097/00008571-200103000-00008;
 RA Ito S., Ito I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;
 RT "Polymorphism of the ABC transporter genes, MRP1, MRP2 and MRP3/ABCC1,
 in healthy Japanese subjects.";
 RL Pharmacogenetics 11:175-184 (2001).
 CC -1- FUNCTION: May participate directly in the active transport of
 drugs into subcellular organelles or influence drug distribution
 indirectly.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=8;
 CC Comment=Additional isoforms seem to exist. Experimental
 CC confirmation may be lacking for some isoforms;
 CC Name=All isoforms;
 CC IsoId=P33527-1; Sequence=Displayed;
 CC Name=Delexon-17;
 CC IsoId=P33527-2; Sequence=VSP_000037;
 CC Name=Delexon-18;
 CC IsoId=P33527-3; Sequence=VSP_000038;
 CC Name=Delexon-30;
 CC IsoId=P33527-4; Sequence=VSP_000039;
 CC Name=Delexon-17-18;
 CC IsoId=P33527-5; Sequence=VSP_000038;
 CC Name=Delexon-17-30;
 CC IsoId=P33527-6; Sequence=VSP_000037, VSP_000039;
 CC Name=Delexon-18-30;
 CC IsoId=P33527-7; Sequence=VSP_000038, VSP_000039;
 CC Name=Delexon-17-18-30;
 CC IsoId=P33527-8; Sequence=VSP_000037, VSP_000039;
 CC -1- TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear
 CC cells.
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
 CC -1- SIMILARITY: Contains 2 ABC transporter domains.
 CC -1- DATABASIS: NME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiochem.fr/services/chromocancer/genes/MRP106.html".
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; L05628; AAB4616.1; -, mRNA.
 CC EMBL; AF022853; AAB83983.1; -, Genomic DNA.
 CC EMBL; AF022854; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022825; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022826; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022827; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022828; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022829; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022830; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022831; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022832; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022833; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022834; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022835; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022836; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022837; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022838; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022839; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022840; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022841; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022842; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022843; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022844; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022845; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022846; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022847; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL; AF022848; AAB83983.1; JOINED; Genomic DNA.

DR EMBL; AF022849; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83983.1; JOINED; Genomic DNA.

Query Match 79.4%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALRGFCADGSDPLMDNMTNMTSNPDKPCONTLVVWPCGYLACGPFELYSRH 60
DB 1 MALRGFCADGSDPLMDNMTNMTSNPDKPCONTLVVWPCGYLACGPFELYSRH 60
QY 61 DRGYIQMTPLNKTKTALGFLMTIVCMADLFYSFWEBSRGIFLAPVFLVSTLIGITLLA 120
DB 61 DRGYIQMTPLNKTKTALGFLMTIVCMADLFYSFWEBSRGIFLAPVFLVSTLIGITLLA 120
QY 121 TELIOERRKGVOSSGIMLTFMLVALCALAIIRSKIMTALKEDAQVDLFRDITFYVFS 180
DB 121 TELIOERRKGVOSSGIMLTFMLVALCALAIIRSKIMTALKEDAQVDLFRDITFYVFS 180
QY 181 LLLIOLVLSGFSRSPFSETHDNPCESSASFISRIEFTWITGLIYRGYQPLEGSD 240
DB 181 LLLIOLVLSGFSRSPFSETHDNPCESSASFISRIEFTWITGLIYRGYQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKYVSSKDPQPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKYVSSKDPQPKSSKYDANEVEAL 300
QY 301 IVKSPKKNPSPFLVLYKTFGPFYFLMSFFPKAIDIMLMSGPOILIKLVNDTKAPD 360
DB 301 IVKSPKKNPSPFLVLYKTFGPFYFLMSFFPKAIDIMLMSGPOILIKLVNDTKAPD 360
QY 361 MOGFYTVLLFVTAQOTLVHOFHICFVSGMRITKAVIGAYYRKALVITNSARKSTV 420
DB 361 MOGFYTVLLFVTAQOTLVHOFHICFVSGMRITKAVIGAYYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLLMTLNGPSVLAGAVALWMPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLLMTLNGPSVLAGAVALWMPVN 480
QY 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILINGIIVLKLIVAMELAFKDKVLAIRQEBLVK 540
DB 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILINGIIVLKLIVAMELAFKDKVLAIRQEBLVK 540
QY 541 KSAVLSAVGTFVTCPELVALCTPAVYVITDENNIIIDACTAVSLAFNIIARPLNIP 600
DB 541 KSAVLSAVGTFVTCPELVALCTPAVYVITDENNIIIDACTAVSLAFNIIARPLNIP 600
QY 601 MTSSIVQASVSLKRLRIFLSHELEBPSIERRPVKDGGGTSITVANAFTWARSDDPT 660
DB 601 MTSSIVQASVSLKRLRIFLSHELEBPSIERRPVKDGGGTSITVANAFTWARSDDPT 660
QY 661 LINGITFISPEGALVAVVGQVCGKSSILSALLAEMDKVEGHVAIKGSVAVPOAMIOND 720
DB 661 LINGITFISPEGALVAVVGQVCGKSSILSALLAEMDKVEGHVAIKGSVAVPOAMIOND 720
QY 721 SLRENIILFCQLEEPYRSYIOACALLPDLFIIPSGDRTIEGKCVNLSCGQKORSILAR 780
DB 721 SLRENIILFCQLEEPYRSYIOACALLPDLFIIPSGDRTIEGKCVNLSCGQKORSILAR 780
QY 781 AAVSNADIYLFDDPLSAVDAHVGHKIPENYIGPKMKDKNTRILVTHSMYSILPOVDYIV 840
DB 781 AAVSNADIYLFDDPLSAVDAHVGHKIPENYIGPKMKDKNTRILVTHSMYSILPOVDYIV 840
QY 841 MSGGKISEMGSYOELLARDGAFAPLRTYASTBOEDAEENGTVGSGPKAKOMENG 900
DB 841 MSGGKISEMGSYOELLARDGAFAPLRTYASTBOEDAEENGTVGSGPKAKOMENG 900
QY 901 LVYDSAGKOLQORLSSSSYSGDIISRHNTAFIQRAKCKERTYKLMEDKXQOTGVKL 960
DB 901 LVYDSAGKOLQORLSSSSYSGDIISRHNTAFIQRAKCKERTYKLMEDKXQOTGVKL 960
QY 961 SVYWDYKAIKGLFISFLIFMCMNVSAALASNYMLSLMTDDEIVNGTOHTKVRLSVYG 1020
DB 961 SVYWDYKAIKGLFISFLIFMCMNVSAALASNYMLSLMTDDEIVNGTOHTKVRLSVYG 1020

```

```

DB 961 SVYWDYKAIKGLFISFLIFMCMNVSAALASNYMLSLMTDDEIVNGTOHTKVRLSVYG 1020
QY 1021 ALGISGIAVFGSMASVIGGIIASRCILHYDLHSLIRSPMSFEETPSGNLVNRSKEL 1080
DB 1021 ALGISGIAVFGSMASVIGGIIASRCILHYDLHSLIRSPMSFEETPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPEVIYKMFMSLFENVIGACIVILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
DB 1081 DTVDSMIPEVIYKMFMSLFENVIGACIVILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
QY 1141 KRLESVSRSPVSHFNETLGVSVIRAFEBQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
DB 1141 KRLESVSRSPVSHFNETLGVSVIRAFEBQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
QY 1201 VRLCYGNCIVLPAALPAVYSRSLAGLVGSLVSVLSQVTTYLNLVYKSSSEMETRYVA 1260
DB 1201 VRLCYGNCIVLPAALPAVYSRSLAGLVGSLVSVLSQVTTYLNLVYKSSSEMETRYVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSMPQVGRVGFERNYCLRYAREDDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSMPQVGRVGFERNYCLRYAREDDLPVLRHINVTINGG 1320
QY 1321 EKVGIYGRGAGKSSVTLGLFRINESAGEIIIDGINAKIGHDLRFKIIIPDPVLF 1380
DB 1321 EKVGIYGRGAGKSSVTLGLFRINESAGEIIIDGINAKIGHDLRFKIIIPDPVLF 1380
QY 1381 SGLRNLNLPFSQYSDPEWTSLELAHLKDFVSLPDKLDHCEAGGENISVQOROLVCL 1440
DB 1381 SGLRNLNLPFSQYSDPEWTSLELAHLKDFVSLPDKLDHCEAGGENISVQOROLVCL 1440
QY 1441 ARALKRKTILVDEAPVADLETDDLIQSTIRTOPEDDCVLTIAHRLNTIMTYVIVL 1500
DB 1441 ARALKRKTILVDEAPVADLETDDLIQSTIRTOPEDDCVLTIAHRLNTIMTYVIVL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531

RESULT 2
Q9U099 HUMAN
ID Q9U099 HUMAN PRELIMINARY; PRT; 1515 AA.
AC Q9U099
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.B., Kurz E.U., Cole S.P.C., Deley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DB EMBL; AF022853; AAB83979.1; -, Genomic DNA.
DB EMBL; AF022827; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022828; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022829; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022831; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022833; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022835; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022837; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022839; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022841; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022850; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022849; AAB83979.1; JOINED; Genomic DNA.

```

DR EMBL; AF022848; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022840; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022838; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022824; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022823; AAB83979.1; JOINED; Genomic DNA.
DR HSP; F08716; 1MT0.
DR Ensembl; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity; coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC TM transport.
DR InterPro; IPR003439; ABC_transport_like.
DR InterPro; IPR005292; NRP_assoc.
DR InterPro; IPR007192; Prot_kinase.
DR Pfam; PF00664; ABC_tran; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR Prodom; PD00006; ABC_transporter; 2.
DR SMART; SM00382; AAA_2.
DR TIGRFAMs; TIGR00957; NRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR ATP-binding; Nucleotide-binding.
DR NON_TER 1 1
SQ SEQUENCE 1515 AA; 169853 MW; DC95592817C439FE CRC64;
Query Match 78.5%; Score 7769; DB 2; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 377 QTLVHGYFHCYSGMRITAVIGAVYRKALVITNSARKSSVYGEIVNLSVDQRFMD 436
DB 361 QTLVHGYFHCYSGMRITAVIGAVYRKALVITNSARKSSVYGEIVNLSVDQRFMD 420
QY 437 LATYINWISAPQVITATVLTMLNIGPSVLAGAVVWVWPVNAVWAMTKTYQVAMHK 496
DB 421 LATYINWISAPQVITATVLTMLNIGPSVLAGAVVWVWPVNAVWAMTKTYQVAMHK 480
QY 497 SKNRIKLAMEIILNGIKVLKLYAMELAFKDKVLAIROBEIKVLKKSAYLSAVGTFTVCT 556
DB 481 SKNRIKLAMEIILNGIKVLKLYAMELAFKDKVLAIROBEIKVLKKSAYLSAVGTFTVCT 540
QY 557 PELVALCTPVAVYTTIDENNITLDAQFAVSALFNILFPPLNIPMVYSSIVQSVSLKRL 616
DB 541 PELVALCTPVAVYTTIDENNITLDAQFAVSALFNILFPPLNIPMVYSSIVQSVSLKRL 600
QY 617 RIFLSHEELPDSIERPVDGSGTNSITVRNATFTWARSDDPTLNGITFSIPGALVAV 676
DB 601 RIFLSHEELPDSIERPVDGSGTNSITVRNATFTWARSDDPTLNGITFSIPGALVAV 660
QY 677 VGVGCGKSSILSALTAEKDKVEGHVAIKGSVAVVPOQAWIIONDSLRENTLFGQLEBPY 736
DB 661 VGVGCGKSSILSALTAEKDKVEGHVAIKGSVAVVPOQAWIIONDSLRENTLFGQLEBPY 720
QY 737 YRSVYQACALLPDLEIILPSGDRTEIGEKVYNSGGQKQVSLARAVYSNADYILFDPLS 796
DB 721 YRSVYQACALLPDLEIILPSGDRTEIGEKVYNSGGQKQVSLARAVYSNADYILFDPLS 780
QY 797 AVDAHVGKHLFENVIGBKGKMLKNKTRILVYHSYSYPOVUVIIVMSGGKISEMSYQEL 856
DB 781 AVDAHVGKHLFENVIGBKGKMLKNKTRILVYHSYSYPOVUVIIVMSGGKISEMSYQEL 840
QY 857 ARDGAFAEFLRTYASTEOEDAEENGVTGSGGKEAKQKQENGLVYDSAGKOLORLSS 916
DB 841 ARDGAFAEFLRTYASTEOEDAEENGVTGSGGKEAKQKQENGLVYDSAGKOLORLSS 900
QY 917 SSSYSGDISRHNSVTAELQAEAKGETWKLMEADKQCTQVLSVYWDYKAIGLFISF 976
DB 901 SSSYSGDISRHNSVTAELQAEAKGETWKLMEADKQCTQVLSVYWDYKAIGLFISF 960
QY 977 LSTFLPCHNVASLASVYMLSTMTDPIVNGTOHTVRLSVYGALGISGIAVFGSMA 1036
DB 961 LSTFLPCHNVASLASVYMLSTMTDPIVNGTOHTVRLSVYGALGISGIAVFGSMA 1020
QY 1037 VSIGILASRCLAVDLHSLRSFMSFEFRTPSGNTVNRFSKELDYDSMIPVIKAFMG 1096
DB 1021 VSIGILASRCLAVDLHSLRSFMSFEFRTPSGNTVNRFSKELDYDSMIPVIKAFMG 1080
QY 1097 SLENVIGACIVIIILATPIAAIIIPPLGLIYFYVORFYVASSRQIKLESVSRSPYSHFN 1156
DB 1081 SLENVIGACIVIIILATPIAAIIIPPLGLIYFYVORFYVASSRQIKLESVSRSPYSHFN 1140
QY 1157 ETLVGSVIRAFEEQEFTHQSDLKVDENOKAYYPSIVARWMLAVRECVGNCIVLPAAL 1216
DB 1141 ETLVGSVIRAFEEQEFTHQSDLKVDENOKAYYPSIVARWMLAVRECVGNCIVLPAAL 1200
QY 1217 FAVISRSLSAGLVGSLVSLOYTTYLNLVNRSSMEETNIVAVEBLKESSETEKEAPW 1276
DB 1201 FAVISRSLSAGLVGSLVSLOYTTYLNLVNRSSMEETNIVAVEBLKESSETEKEAPW 1260
QY 1277 QIOETAPSSWPQVGRVEFRNYCLARYEDLPVLRHINVTINGEKVIGVGRTAGKSSL 1336
DB 1261 QIOETAPSSWPQVGRVEFRNYCLARYEDLPVLRHINVTINGEKVIGVGRTAGKSSL 1320
QY 1337 TLGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQPVLPFGSLRNMLDPFSQYSD 1396
DB 1321 TLGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQPVLPFGSLRNMLDPFSQYSD 1380
QY 1397 BEVWTSLELHLKDFVSLPDKLDHECAEGENLSVGOQOLVCARALLRKTKLIVDEA 1456
DB 1381 BEVWTSLELHLKDFVSLPDKLDHECAEGENLSVGOQOLVCARALLRKTKLIVDEA 1440
QY 1457 TAAVDETDLLIOSTIRTOFEDCTVTLIAHRLNTIMDYEVIVLKDGEIOEYGAPSDLLQ 1516

Db 1441 TAAVLETDLDLQSTIRTFQEDCTVLTARLNTIMTYRVIYLDKSGIQEYGAPSDLO 1500
 QY 1517 ORGLFYSMADAGLV 1531
 Db 1501 ORGLFYSMADAGLV 1515

RESULT 3

08649 MACFA
 ID 08649 MACFA PRELIMINARY; PRT; 1531 AA.
 AC 08649;
 DT 01-JUN-2003 (TREMBlurel. 24, Created)
 DT 01-JUN-2003 (TREMBlurel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlurel. 26, Last annotation update)
 DE Multidrug resistance protein 1B.
 GN Name=MRP1;
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopithecoidea; Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP NOCLEOTIDE SEQUENCE.
 RX MEDLINE=22544876; PubMed=12657726;
 RA Godinot N., Iversen P.W., Tabes L., Xia X., Williams D.C.,
 RA Dantzig A.H., Perry W.L.,
 RT "Cloning and functional characterization of the multidrug resistance-
 RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
 RL Mol. Cancer Ther. 2:307-316(2003).
 RN [2]
 RP NOCLEOTIDE SEQUENCE.

RA Perry W.L. III, Godinot N.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; AY146673; AAN65349.1; -; mRNA.
 DR HSSP; P08716; 1MT0.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016887; F:ATPase activity; IEA.
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transporter; IEA.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR011527; ABC_Membrane_1.
 DR InterPro; IPR001140; ABC_TM_transp.
 DR InterPro; IPR003439; ABC_transp_like.
 DR InterPro; IPR005292; MRP_assoc.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE; PS50929; ABC_TMIF; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 KW ATP-binding; Nucleotide-binding; Repeat.
 SQ SEQUENCE 1531 AA; 171659 MW; 1AE788BDF9EF459 CRC64;

Query Match 77.7%; Score 7695; DB 2; Length 1531;
 Best local Similarity 97.5%; Pred. No. 0;
 Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTANTSNDFTRKCFONTVTVVWPCFYLMACFPFYFLYSRH 60
 Db 1 MALRGFCSADGSDPLMDMNTANTSNDFTRKCFONTVTVVWPCFYLMACFPFYFLYSRH 60
 QY 61 DRGVIQMTLNTKTKTALGFLMTIVCWADLFYSFWERSRGIFLAPVFLVSFTLLGITTLA 120
 Db 61 DRGVIQMTLNTKTKTALGFLMTIVCWADLFYSFWERSRGIFLAPVFLVSFTLLGITTLA 120

QY 121 TFLIQERRKGVSSGIMLTFWLVALYCALAILRSKIMTALKEDAQVDFRDIIFYYFS 180
 Db 121 TFLIQERRKGVSSGIMLTFWLVALYCALAILRSKIMTALKEDAQVDFRDIIFYYFS 180
 QY 181 LLLIQLVLSGFSRDRSPLFSETIHDNPNCPSSASFLSRITFWMTGLIVRGYRPLGSD 240
 Db 181 LVLIIQLVLSGFSRDRSPLFSETIHDNPNCPSSASFLSRITFWMTGLIVRGYRPLGSD 240
 QY 241 LMSLNKEDTSBOVVPVTVKMKKECATRQKQPVVVVSSQDPAQPKSSKVDANEVEAL 300
 Db 241 LMSLNKEDTSBOVVPVTVKMKKECATRQKQPVVVVSSQDPAQPKSSKVDANEVEAL 300
 QY 301 IVKSPQKKNPSPKPVLYKTFEGPYFLMSFFPKAIHDMFMSSGPOILLIKLKFVNDTAPD 360
 Db 301 IVKSPQKKNPSPKPVLYKTFEGPYFLMSFFPKAIHDMFMSSGPOILLIKLKFVNDTAPD 360
 QY 361 MGGFYTVLLFVYACLOTVLHQYFHCIFVSGRIKTAIVGAVRKALVITNSARKSSTV 420
 Db 361 MGGFYTVLLFVYACLOTVLHQYFHCIFVSGRIKTAIVGAVRKALVITNSARKSSTV 420
 QY 421 GEIYNLMSVDAQRPMDLATTYNNIMWAPLOVILALYLLMNLGFSVLGAVAVMLVPVN 480
 Db 421 GEIYNLMSVDAQRPMDLATTYNNIMWAPLOVILALYLLMNLGFSVLGAVAVMLVPVN 480
 QY 481 AVNANKTKTYQVAHMSKDNRIKLMSLILNGIKVLKLYAMELAFKDVLAIRGELVYK 540
 Db 481 AVNANKTKTYQVAHMSKDNRIKLMSLILNGIKVLKLYAMELAFKDVLAIRGELVYK 540
 QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITIDENNILDAOTAFVSLALFNILRPPLNLP 600
 Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITIDENNILDAOTAFVSLALFNILRPPLNLP 600
 QY 601 MVISSIVQASVSLKRLRIFLSHBLBDSIERRPVKDGGSGTNSITVNNATFTWARSDDPT 660
 Db 601 MVISSIVQASVSLKRLRIFLSHBLBDSIERRPVKDGGSGTNSITVNNATFTWARSDDPT 660
 QY 661 LMGITTSIPRGALVAVVGQVCGKSSLSLALBMDVBEHVALKSGVAVVPQAMVOND 720
 Db 661 LMGITTSIPRGALVAVVGQVCGKSSLSLALBMDVBEHVALKSGVAVVPQAMVOND 720
 QY 721 SLRENILFGQLEEPYRYSVIOACALLPDLIELPSGDRTEIGRQVNLSGGQKORVSLAR 780
 Db 721 SLRENILFGQLEEPYRYSVIOACALLPDLIELPSGDRTEIGRQVNLSGGQKORVSLAR 780
 QY 781 AVYSNADITYPDDPLSAVDHVGKHIPENVIGPKMLKNKTRILVTHSMSYLPQVDYIIV 840
 Db 781 AVYSNADITYPDDPLSAVDHVGKHIPENVIGPKMLKNKTRILVTHSMSYLPQVDYIIV 840
 QY 841 MSGGKISEMGSSYQELLARDGAFELRTYASAEQEDPEDNGVTVGSGPKKAKQENGM 900
 Db 841 MSGGKISEMGSSYQELLARDGAFELRTYASAEQEDPEDNGVTVGSGPKKAKQENGM 900
 QY 901 LVTDSAGKOLQROLSSSSSYSGDISRHHNSTALEOKAEKKEFTKLMLEADKAQOTGVKL 960
 Db 901 LVTDSAGKOLQROLSSSSSYSGDISRHHNSTALEOKAEKKEFTKLMLEADKAQOTGVKL 960
 QY 961 SVYWDYMKALGIFLSFLSIFLPMCNVVSALASNYMLSLMTDDPIVNGTOGHTKVRLSVYG 1020
 Db 961 SVYWDYMKALGIFLSFLSIFLPMCNVVSALASNYMLSLMTDDPIVNGTOGHTKVRLSVYG 1020
 QY 1021 ALGISGIAVFGSMASVSTGILIASRCLHVDLHSLRSMSFFERTPSGNLVNRSKEL 1080
 Db 1021 ALGISGIAVFGSMASVSTGILIASRCLHVDLHSLRSMSFFERTPSGNLVNRSKEL 1080
 QY 1081 DTVDSMIPEVIKPMFMSLFVIGACIVILLATPIAIIIPPLGLIYEFVORFYVASRQL 1140
 Db 1081 DTVDSMIPEVIKPMFMSLFVIGACIVILLATPIAIIIPPLGLIYEFVORFYVASRQL 1140
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAAYPSIVANRWLA 1200
 Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAAYPSIVANRWLA 1200
 QY 1201 VRLCEGVNCIVLPAALFAVVISRHSLSAGLVGSLVSYSLOVTTYTILNLMVMSSEMEFNIVA 1260

Db 1201 VRLECVNAGNCVTLFALPAVLSRHSLSAGLVLSVSYSLQVTTYTNMVLWRMSSEMETIVA 1260
QY 1261 VERLKESETEKEAPMOIOETAPSSMPQVGRVEPRNYCLRYREDLPVLRHINVTNGG 1320
Db 1261 VERLKESETEKEAPMOIOETAPSSMPQVGRVEPRNYCLRYREDLPVLRHINVTNGG 1320
QY 1321 EKVGVIGRTGAGKSLTGLFRINESAGEIIGININAKIGLHDLRKKITIIIPQDPVLF 1380
Db 1321 EKVGVIGRTGAGKSLTGLFRINESAGEIIGININAKIGLHDLRKKITIIIPQDPVLF 1380
QY 1381 SGSRFNMULDPSQYSDEEWTSLSLAHKDFVSALPDLDEHCAEGGENLSVGORQVCL 1440
Db 1381 SGSRFNMULDPSQYSDEEWTSLSLAHKDFVSALPDLDEHCAEGGENLSVGORQVCL 1440
QY 1441 ARALLRKTKIIVLEBATAVLETDLQSTIRFOFEDCTVLTIAHRLNTIMDYTRVVL 1500
Db 1441 ARALLRKTKIIVLEBATAVLETDLQSTIRFOFEDCTVLTIAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYMAKQAGLV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYMAKQAGLV 1531

RESULT 4
Q864S0_MACFA PRELIMINARY; PRT; 1531 AA.
ID Q864S0;
AC Q864S0;
DT 01-JUN-2003 (TREMblrel. 24, Created)
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Multidrug resistance protein 1A.
GN Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Gymnorigus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecoinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22544876; PubMed=12657726;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Danczig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Perry W.L., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY146672; AAN65348.1; -; mRNA.
DR HSSP; P08716; IMT0.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD00006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.

DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS0107; PROTEIN KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1531 AA; 171701 MW; 168712B8CC2D2B89 CRC64;
Query Match 77.7%; Score 7689; DB 2; Length 1531;
Best local similarity 97.6%; Pred. No. 0;
Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;
QY 1 MALRGPSAGSDPDLMDMTNTMTNSNDPTKCPONTLVVWPCGYLMACPEFYLYSRH 60
Db 1 MALRGPSAGSDPDLMDMTNTMTNSNDPTKCPONTLVVWPCGYLMACPEFYLYSRH 60
QY 61 DRGYQMTPLNKTKTALGFLLMTCWADLFYSPERSRGFLFAPVFLVSPFLIGITTLTA 120
Db 61 DRGYQMTPLNKTKTALGFLLMTCWADLFYSPERSRGFLFAPVFLVSPFLIGITTLTA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVALCALAIRSKIMTALKEDAVDLFRDITFYVYS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVALCALAIRSKIMTALKEDAVDLFRDITFYVYS 180
QY 121 TFLIQLERRKGVSSGIMLTFWLVALCALAIRSKIMTALKEDAVDLFRDITFYVYS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVALCALAIRSKIMTALKEDAVDLFRDITFYVYS 180
QY 181 LLLIQVLSCGSDPSPLFSETHDNPCESSASFLRITFWMTTGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLSCGSDPSPLFSETHDNPCESSASFLRITFWMTTGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLYKMKKSCAKTRKQPVVYVSSKDPAQPKSSKVDANEYEAL 300
Db 241 LMSLNKEDTSEQVVPVLYKMKKSCAKTRKQPVVYVSSKDPAQPKSSKVDANEYEAL 300
QY 301 IVKSPQKEMNPSEKVLKTFGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTYAPD 360
Db 301 IVKSPQKEMNPSEKVLKTFGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTYAPD 360
QY 301 IVKSPQKEMNPSEKVLKTFGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTYAPD 360
Db 301 IVKSPQKEMNPSEKVLKTFGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTYAPD 360
QY 361 WQGFYFVLLFVACLOTIVLHQYFHCIVSGMRKTAIVGAVYRKALVTNARKSTV 420
Db 361 WQGFYFVLLFVACLOTIVLHQYFHCIVSGMRKTAIVGAVYRKALVTNARKSTV 420
QY 421 GEIVNLSVDAQREMDLATYINMIMSAPLOVILALYLLMNTIGPSVLAVAVVLMVFN 480
Db 421 GEIVNLSVDAQREMDLATYINMIMSAPLOVILALYLLMNTIGPSVLAVAVVLMVFN 480
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLYLWELAKDYALAIROBELKVLK 540
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLYLWELAKDYALAIROBELKVLK 540
QY 541 KSAFLSAGFTTWTCTFFVALCTFAVYVITDENIIDAQTAFLSLFNTLRPNLILP 600
Db 541 KSAFLSAGFTTWTCTFFVALCTFAVYVITDENIIDAQTAFLSLFNTLRPNLILP 600
QY 601 MVISIYQASVSLKRLIFLSHELEBPSIERRPVKDGGGNTSTVBNATFTMARSDPT 660
Db 601 MVISIYQASVSLKRLIFLSHELEBPSIERRPVKDGGGNTSTVBNATFTMARSDPT 660
QY 661 LMGITFIPREGALVAVVGVGCGKSLLSALLAEMDRVEGHVAKGSVAVYPQQAMTQND 720
Db 661 LMGITFIPREGALVAVVGVGCGKSLLSALLAEMDRVEGHVAKGSVAVYPQQAMTQND 720
QY 721 SLRNNILFGCOLLEPPYRSVYQACALLPDLEILPSGRTIEGEGVNLSSGQKRVSLAR 780
Db 721 SLRNNILFGCOLLEPPYRSVYQACALLPDLEILPSGRTIEGEGVNLSSGQKRVSLAR 780
QY 781 AVYSNADITYFPDPLSAVDHVGHIPEVNIYGPQMLKNKTRILVTHSMSTLPQVDYIIV 840
Db 781 AVYSNADITYFPDPLSAVDHVGHIPEVNIYGPQMLKNKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGGKISEMSYQELIARDGAFAEFLRTYASTEQDAEENGVTGVSQPKAEAYOMENGM 900
Db 841 MSGGKISEMSYQELIARDGAFAEFLRTYASTEQDAEENGVTGVSQPKAEAYOMENGM 900
QY 901 LVYDSAGKOLOROLSSSSYSISGDI SRHNSYAEIQAQBAKKEEFYKLMKADKATQGVYKL 960
Db 901 LVYDSAGKOLOROLSSSSYSISGDI SRHNSYAEIQAQBAKKEEFYKLMKADKATQGVYKL 960

Oy		961	SVMYDYMKAIGLFIFFLSIFPEMCHVSAKALSNWYLMTWDPTINGQOETHKYRLSYG	1020		
Dd		961	SVYMDYMAIKGLFIFFLSIFLCNHWAAALNSNYLSMTDDPIYNGOETHKYRLSYG	1020		
Oy		1021	ALGISOGIAVFEGYSMAVSIGGIILASCRCHVDLLHSILKS PMSPFERTPSGNLVNRFSKET.	1080		
Dd		1021	ALGISOGIAVFEGYSMAVSIGGIILASRYLHVLDLHSLTSPMSPFERTPSGNLVNRFSKET.	1080		
Oy		1081	DTVDSMIPEVLKMFNGSLPNVI GACIVILLATPILAIITIPPLGLIYFVGQFPYASSSQQL	1140		
Dd		1081	DTVDSMIPEVLKMFNGSLFNVI GACIVILLATPILAIITIPPLGLIYFVGQFPYASSSQQL	1140		
Oy		1141	KRLSEVSRSPPYXSHNETTLIGSVYRAFEEOERFIHQSDLKVDENOKAYPSIVANRWLA	1200		
Dd		1141	KRLSEVSRSPPYXSHNETTLIGSVYRAFEEOERFIHQSDLKVDENOKAYPSIVANRWLA	1200		
Oy		1201	VRLBECVNCIVLPALPAVISRHSISAGLVGLSVSLSQVTYYLNMLVRMSEMETNIIVA	1260		
Dd		1201	VRLBECVNCIVLPALPAVISRHSISAGLVGLSVSLSQVTYYLNMLVRMSEMETNIIVA	1260		
Oy		1261	VERLKEYESETEEAEMWOIQETAPPSNMQVGRVEFRANCIFYREBDLPVLRHINVTINGG	1320		
Dd		1261	VERLKEYESETEEAEMWOIQETAPPSNMQVGRVEFRANCIFYREBDLPVLRHINVTINGG	1320		
Oy		1321	EKVGVVGTGTGAGSKSLTGLFRINSAGEEIIIDGINAKIGHDLRKITITIPODPLYF	1380		
Dd		1321	EKVGVVGTGTGAGSKSLTGLFRINSAGEEIIIDGINARLGHLDRFKITITIPODPLYF	1380		
Oy		1381	SGSLRMNLDPFSQYSDEEWNTSELBAHLKD PFVSALPDKLDRBCABGGENLSVGQROLVCL	1440		
Dd		1381	SGSLRMNLDPFSQYSDEEWNTSELBAHLKG FVSALPDKLDRBCABGGENLSVGQROLVCL	1440		
Oy		1441	ARALLRKTKIIVLDBATAAVULEFTDDLIOSTIRIQEFECTVLTIAHRANTIMDYTRYIVL	1500		
Dd		1441	ARALLRKTKIIVLEDBATAVULEFTDDLIOSTIRIQEFECTVLTIAHRINTIMDYTRYIVL	1500		
Oy		1501	DKGEIOEGYGPSDLLQQRGLFYSSAKDAKGLV	1531		
Dd		1501	DKGEIOEGYGPSDLLQQRGLFYNNARKADGLV	1531		
RESULT 5						
O9UC97 HUMAN						
ID	O9UC97	HUMAN PRELIMINARY;	PRT;	1459 AA.		
AC	O9UC97					
DT	01-MAY-2000	(TREMBLrel. 13, Created)				
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)				
DT	01-MAR-2004	(TREMBLrel. 26, Last annotation update)				
DE	Multidrug resistance protein (Fragment).					
GN	Name=MRP;					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;					
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;					
OC	Homo.					
OX	NCHI_1	Taxid=9606;				
RN						
RP	NUCLEOTIDE SEQUENCE.					
EX	MEDLINE=96008927;	PubMed=9344662;	DOI=10.1006/geno.1997.4950;			
FA	Grant C.E., Kurz E.-U., Cole S.P.C., Deeley R.G.;					
RT	"Analysis of the intron-exon organization of the human multidrug-					
RL	resistance protein gene (MRP) and alternative splicing of its mRNA.";					
CC	Genomics 45:368-378(1997).					
-	SIMILARITY: Belongs to the ABC transporter family.					
EMBL	AF022853;	AAB83980.1	-- Genomic DNA.			
DR	EMBL	AF022824;	AAB83980.1	JOINED; Genomic DNA.		
DR	EMBL	AF022826;	AAB83980.1	JOINED; Genomic DNA.		
DR	EMBL	AF022828;	AAB83980.1	JOINED; Genomic DNA.		
DR	EMBL	AF022830;	AAB83980.1	JOINED; Genomic DNA.		
DR	EMBL	AF022832;	AAB83980.1	JOINED; Genomic DNA.		
DR	EMBL	AF022834;	AAB83980.1	JOINED; Genomic DNA.		
DR	EMBL	AF022836;	AAB83980.1	JOINED; Genomic DNA.		

	DR	EMBL; AF022838; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022848; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022847; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022846; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022845; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022844; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022843; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022842; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022841; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022839; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022838; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022837; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022836; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022835; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022834; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022833; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022832; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022831; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022830; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022829; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022828; AAB83980.1; JOINED; Genomic DNA.
	DR	HSSP; P08716; LMT0.
	DR	Ensembl; ENSG00000103222; Homo sapiens.
	DR	GO; GO:0016021; C:integral to membrane; IEA.
	DR	GO; GO:0005524; F:ATP binding; IEA.
	DR	GO; GO:0016887; F:ATPase activity; IEA.
	DR	GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
	DR	GO; GO:0000166; F:nucleotide binding; IEA.
	DR	GO; GO:0006810; P:transport; IEA.
	DR	InterPro; IPR003593; AAA ATPase.
	DR	InterPro; IPR011527; ABC_membrane_1.
	DR	InterPro; IPR001140; ABC_TM_transp.
	DR	InterPro; IPR003439; ABC_transp_lkte.
	DR	InterPro; IPR005292; NRP_assoc.
	DR	InterPro; IPR000719; Prot_kinase.
	DR	Pfam; PF00664; ABC_membrane; 2.
	DR	Pfam; PF00005; ABC_tran; 1.
	DR	Prodrom; PD000006; ABC_transporter; 1.
	DR	SMART; SM00382; AAA; 2.
	DR	TIGRFAMs; TIGR00957; NRP_assoc_pro; 1.
	DR	PROSITE; PS00929; ABC_TMlf; 2.
	DR	PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
	DR	PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
	DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
	KW	ATP-binding; Nucleotide-binding.
	FT	NON TER
	SQ	SEQUENCE 1459 AA; 163831 MW; 6A016A5AE7AA95C CRC64;
	Query Match	75.2%; Score 7444; DB 2; Length 1459;
	Best Local Similarity	96.3%; Pred. No. 0;
	Matches 1459; Conservative	0; Mismatches 0; Indels 56; Gaps 1;
OY	17	DNNVTNNTSNDPDFKFCQNTVTVVWPCCFYLMACPFYLYTSRRDRGTOWTPIINKTKTA 76
Db	1	DMNVTNNTSNDPFKFCQNTVTVVWPCCFYLMACPFPFLYSRRDRGTOWTPIINKTKTA 60
OY	77	LGFLLMTIVCMADLTFYSFWEBSRGIFLAIPVLVSFTLLGITTLATFTLIQLERRKGVOSSG 136
Db	61	LGFLLMTIVCMADLTFYSFWERSRGIFLAIPVLVSFTLLGITTLATFTLIQLERRKGVOSSG 120
OY	137	IMLTFFMVALVCALAIIRSKIMTALKEDAQVDLFREDITFYVYFSLLLIQLVLSCPSDRSP 196
Db	121	IMLTFFMVALVCALAIIRSKIMTALKEDAQVDLFREDITFYVYFSLLLIQLVLSCPSDRSP 180
OY	197	LFSETHIDPNCPRESSASPLSRITFMWTITGLIYNVGROPLEGSLMLNKEDTSBOVVAV 256
Db	181	LFSETHIDPNCPRESSASPLSRITFMWTITGLIYNVGROPLEGSMLSNKEDTSBOVVAV 240
OY	257	LVKNMKKECATRIKQPVKVYVSSKDPAOPKSSSKVDANEEVEALIVSPKEMNPISLFKV 316
Db	241	LVKNMKKECATRIKQPVKVYVSSKDPAOPKSSSKVDANEEVEALIVSPKEMNPISLFKV 300
OY	317	LYKTGGPYFLMSPPFKAIHDLNMFSGQILKLIIKYVNDITAPDWOGYFTYVLLPVTAQL 376

Db 301 LYKTEGPEYFLMSFEFKAIHDLMMFSGPQIILKLIKFVNDTKAPDMOGYFTYVLLFVTAQL 360
Qy 377 QTLVHOYFHI CFSGNRIKTAIVIGAYRKALVITNSAKSSTGEIYNLMSVDAQSFMD 436
Db 361 QTLVHOYFHI CFSGNRIKTAIVIGAYRKALVITNSAKSSTGEIYNLMSVDAQSFMD 420
Qy 437 LATYNNIWSAPLQVILALYLLMLNGPSVLAGVAVMVLVAVNAVMAKTKTYQVAHMK 496
Db 421 LATYNNIWSAPLQVILALYLLMLNGPSVLAGVAVMVLVAVNAVMAKTKTYQVAHMK 480
Qy 497 SKDNRIKLMEEIINGIKVLYKLYAMELAFKDYLAIRQELKVKLKSAYLSVGFYTWCT 556
Db 481 SKDNRIKLMEEIINGIKVLYKLYAMELAFKDYLAIRQELKVKLKSAYLSVGFYTWCT 540
Qy 557 PFLVALCTFAYVYTIIDENNII DAOAFYSALFNILRPILNIPVNISSIQASVSLKRL 616
Db 541 PFLVALCTFAYVYTIIDENNII DAOAFYSALFNILRPILNIPVNISSIQASVSLKRL 600
Qy 617 RIFLSHEELEPDSIERRPVKDGGTNSITVNNATFTMARSDPPTLNGITFSIPGALVAV 676
Db 601 RIFLSHEELEPDSIERRPVKDGGTNSITVNNATFTMARSDPPTLNGITFSIPGALVAV 660
Qy 677 VGQVCGCKSSLLSLALAMDVEGHVAIKGSVAVVPQAMVQNDISRENIILFGQLEBPY 736
Db 661 VGQVCGCKSSLLSLALAMDVEGHVAIKGSVAVVPQAMVQNDISRENIILFGQLEBPY 720
Qy 737 YRSYIOACALLPDEILPESGRTEIGEIKVNLSSGQKQVSLARAVYSNADIVLFDPLS 796
Db 721 YRSYIOACALLPDEILPESGRTEIGE----- 747
Qy 797 AVDAHVGKHIFENYIGPKGMLKNKTRII VTHSMSTYLPQVDYIIVMSGKISEMSGYOELL 856
Db 748 -----KTRILVTHSMSTYLPQVDYIIVMSGKISEMSGYOELL 784
Qy 857 ARDQAPAFELRTYASTEBODAEENGVTGVSGPKGEAQMENGMLVYTSASQKQOROLSS 916
Db 785 ARDQAPAFELRTYASTEBODAEENGVTGVSGPKGEAQMENGMLVYTSASQKQOROLSS 844
Qy 917 SSSYSGDISRHHNSTAELQKAKEEFTWKLEADKATQGVQKLSVYDYNKAIIGLPISTP 976
Db 845 SSSYSGDISRHHNSTAELQKAKEEFTWKLEADKATQGVQKLSVYDYNKAIIGLPISTP 904
Qy 977 LSIPLFMCNHVSALASNMYLSLMTDDPIVNGTOEHTKVRLSVYGAIGISQGIANVFGSMA 1036
Db 905 LSIPLFMCNHVSALASNMYLSLMTDDPIVNGTOEHTKVRLSVYGAIGISQGIANVFGSMA 964
Qy 1037 VSIQGIILASRCLAHVDLHSHIIRSPMSFFERTPSCGNLVNRFSEKELDTVDSMTPEVIKMGWG 1024
Db 965 VSIQGIILASRCLAHVDLHSHIIRSPMSFFERTPSCGNLVNRFSEKELDTVDSMTPEVIKMGWG 1024
Qy 1097 SLFNVIGACIYIILATPIAAIIPPLGLIYFFVORFFVASSRQKRLSVSRSPYSHFN 1156
Db 1025 SLFNVIGACIYIILATPIAAIIPPLGLIYFFVORFFVASSRQKRLSVSRSPYSHFN 1084
Qy 1157 ETLGVSIVIRAFEEQERFIHOSDLKVDENOKAYYSIVANRWLAARLECVNGCIYLFPAAL 1216
Db 1085 ETLGVSIVIRAFEEQERFIHOSDLKVDENOKAYYSIVANRWLAARLECVNGCIYLFPAAL 1144
Qy 1217 FAVISRHSLSAGLVGLSVSYSLQVTTYLANVLVRNMSSEMETNIIVAEERLKEYSETEKAPW 1276
Db 1145 FAVISRHSLSAGLVGLSVSYSLQVTTYLANVLVRNMSSEMETNIIVAEERLKEYSETEKAPW 1204
Qy 1277 QIOETAPSSMPOVGRVFRNYCLAYREBDLPVLHINVTIINGGKVGIVGTGAGKSSL 1336
Db 1205 QIOETAPSSMPOVGRVFRNYCLAYREBDLPVLHINVTIINGGKVGIVGTGAGKSSL 1264
Qy 1337 TLGIFRINESAGEIIIDGINIAKIGLHDLREKTIIPQDPVLPFGSGLRMLNDPPSOYSD 1396
Db 1265 TLGIFRINESAGEIIIDGINIAKIGLHDLREKTIIPQDPVLPFGSGLRMLNDPPSOYSD 1324
Qy 1397 BEVWTSLELAHKDFVSALPDKLDHCECABGGENISVGOQOLVCLARALLRKTKIIVLDEA 1456
Db 1325 BEVWTSLELAHKDFVSALPDKLDHCECABGGENISVGOQOLVCLARALLRKTKIIVLDEA 1384

Qy 1457 TAAVDETDLLIGSTIRTOPEDCVLTIAHRLNTINDYTRVIVLDKGEIOEGAPSDLLQ 1516
Db 1385 TAAVDETDLLIGSTIRTOPEDCVLTIAHRLNTINDYTRVIVLDKGEIOEGAPSDLLQ 1444
Qy 1517 ORGLFYSMAQDAGLV 1531
Db 1445 ORGLFYSMAQDAGLV 1459

RESULT 6
QY09040 HUMAN
ID Q9040_HUMAN PRELIMINARY, PRT, 1456 AA.
AC Q9040;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
OC Homo.
CX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
RT resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL, AF022853; AAB83981.1; -; Genomic DNA.
DR EMBL, AF022824; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022825; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022826; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022827; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022829; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022831; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022833; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022835; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022837; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022847; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022846; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022845; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022844; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022843; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022842; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022841; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022840; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022838; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022852; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022851; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022850; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022849; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022848; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022836; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022834; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022832; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022830; AAB83981.1; JOINED; Genomic DNA.
DR EMBL, AF022828; AAB83981.1; JOINED; Genomic DNA.
DR HSSP, P08716, 1MT0.
DR Ensembl, ENSG00000103222; Homo sapiens.
DR GO, GO:0016021, C:integral to membrane, IEA.
DR GO, GO:0005524, F:ATP binding, IEA.
DR GO, GO:0016887, F:ATPase activity, IEA.
DR GO, GO:0042626, F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO, GO:0000166, F:nucleotide binding, IEA.
DR GO, GO:0006810, P:transport, IEA.
DR InterPro, IPR003593, AAA_ATPase.
DR InterPro, IPR011527, ABC_membrane_1.
DR InterPro, IPR001140, ABC_TM_transp.
DR InterPro, IPR003439, ABC_cranep_like.

DR InterPro: IPR005292; MRP assoc.
 DR InterPro: IPR000719; Prot. kinase.
 DR Pfam: PF00664; ABC membrane; 2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR Prodom: PD000006; ABC_transporter; 2.
 DR SMART: SM00382; AAA; 2.
 DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE: PS00929; ABC_TMIF_2; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR ATP-binding; Nucleotide-binding.
 FT NON TER 1
 SQ SEQUENCE 1456 AA; 163233 MW; 8DB8AB2BC481F2 CRC64;

Query Match 74.9%; Score 7419.5; DB 2; Length 1456;
 Best Local Similarity 96.1%; Pred. No. 0;
 Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;

QY 17 DANVTMTSNPDFTKCFONTVLVWPCFYLWACFPFYLYSRDRGYIOWTPIKTKTA 76
 DB 1 DANVTMTSNPDFTKCFONTVLVWPCFYLWACFPFYLYSRDRGYIOWTPIKTKTA 60

QY 77 LGFLMWIVCADLFFYSFWEBSRGIFLAPVPLVSPITLIGITTLATFLLQERRKGVSSG 136
 DB 61 LGFLMWIVCADLFFYSFWEBSRGIFLAPVPLVSPITLIGITTLATFLLQERRKGVSSG 120

QY 137 IMTLFMTVALVCAIAIRSKMTALKEBDAQVDLFRDITFYVFSLLIQLVLSCFSDRSP 196
 DB 121 IMTLFMTVALVCAIAIRSKMTALKEBDAQVDLFRDITFYVFSLLIQLVLSCFSDRSP 180

QY 197 LFSETIHDPNCPSSASPLSRITFPMWITGLVARGYOPLEGSDMLSKEDTSEOVVPV 256
 DB 181 LFSETIHDPNCPSSASPLSRITFPMWITGLVARGYOPLEGSDMLSKEDTSEOVVPV 240

QY 257 LVNWKKECAKTRKQPKVYVSSKDPAPKSSKVDANEVEALIVKSPKEMNPISLFKV 316
 DB 241 LVNWKKECAKTRKQPKVYVSSKDPAPKSSKVDANEVEALIVKSPKEMNPISLFKV 300

QY 317 LYKTGGYFLMSFFFKAIHDLMSFGPQILKLIFVNDTKAPDWQGYFTYVILLFVACL 376
 DB 301 LYKTGGYFLMSFFFKAIHDLMSFGPQILKLIFVNDTKAPDWQGYFTYVILLFVACL 360

QY 377 QTLVHOYFHCFSVGRITKAVTGAVYRKALVITNSARKSTVGEIVNLMSVDAGQFMD 436
 DB 361 QTLVHOYFHCFSVGRITKAVTGAVYRKALVITNSARKSTVGEIVNLMSVDAGQFMD 420

QY 437 LATYINMWASAPLOVILALYLLMLNLGPSVLGAVAVWVLMVAVNAWAMTKTKYQVAHMK 496
 DB 421 LATYINMWASAPLOVILALYLLMLNLGPSVLGAVAVWVLMVAVNAWAMTKTKYQVAHMK 480

QY 497 SKONRIKLMEIILNGIKVLKLYAMELAFKDKVLAIRQEBELVKLKKSAVLSAVGTFTWCT 556
 DB 481 SKONRIKLMEIILNGIKVLKLYAMELAFKDKVLAIRQEBELVKLKKSAVLSAVGTFTWCT 540

QY 557 PELVALCTFAVYVYIDNNILDAQTAVALPILILAFPLNLLPMVSSIVQASVSLKRL 616
 DB 541 PELVALCTFAVYVYIDNNILDAQTAVALPILILAFPLNLLPMVSSIVQASVSLKRL 600

QY 617 RIFLSHELEPDSIERRPVKGSGTNSITYRNATFTVARSDPPTLNGITISIPGALVAV 676
 DB 601 RIFLSHELEPDSIERRPVKGSGTNSITYRNATFTVARSDPPTLNGITISIPGALVAV 660

QY 677 VGVGVCKSSULSALLAEMDKVEGHVAIKGSVAVYVPOQAMIQNDLSRENILLFGQLEBPY 736
 DB 661 VGVGVCKSSULSALLAEMDKVEGHVAIKGSVAVYVPOQAMIQNDLSRENILLFGQLEBPY 688

QY 737 YRSVYQACALLPDLLEIPSGDRTEIGKGVNLSSGCKORVSIARAAYVSNADIVLFDDPLS 796
 DB 689 YRSVYQACALLPDLLEIPSGDRTEIGKGVNLSSGCKORVSIARAAYVSNADIVLFDDPLS 721

QY 797 AVDAHVGHKIFENVIGPKGMLKNKTRILVTHSMGYLPQVVIIVMSGKISSEMGSYOELL 856
 DB 797 AVDAHVGHKIFENVIGPKGMLKNKTRILVTHSMGYLPQVVIIVMSGKISSEMGSYOELL 856

DB 722 AVDAHVGHKIFENVIGPKGMLKNKTRILVTHSMGYLPQVVIIVMSGKISSEMGSYOELL 781

QY 857 ARDGAFAEPLRTVASTBOBDAEENGVTGVSGPCKEAKOMENGMVLVYDSAGKQOROLSS 916
 DB 782 ARDGAFAEPLRTVASTBOBDAEENGVTGVSGPCKEAKOMENGMVLVYDSAGKQOROLSS 841

QY 917 SSSYSGDISRHHNSTAELOKAFAKERTWMLKEADKQOTQVKLSVYWDYKAIGLFTSF 976
 DB 842 SSSYSGDISRHHNSTAELOKAFAKERTWMLKEADKQOTQVKLSVYWDYKAIGLFTSF 901

QY 977 LSIFLFCMCHVASALASNYMLSLWTDDEIVNGTOEHTYKRLSVYGALGISQIAVFGYSMA 1036
 DB 902 LSIFLFCMCHVASALASNYMLSLWTDDEIVNGTOEHTYKRLSVYGALGISQIAVFGYSMA 961

QY 1037 VSTIGILASRCILNVDLHSLIRSPMSFEPTPSGNLVNRSKELDTYDSMTPEYIKHPMG 1096
 DB 962 VSTIGILASRCILNVDLHSLIRSPMSFEPTPSGNLVNRSKELDTYDSMTPEYIKHPMG 1021

QY 1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFVVASRQLKLESYSRSPVYSHFN 1156
 DB 1022 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFVVASRQLKLESYSRSPVYSHFN 1081

QY 1157 ETLIGSVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECYGNCTVLPAL 1216
 DB 1082 ETLIGSVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECYGNCTVLPAL 1141

QY 1217 FAVISRHSLSAGLVGSLVSVSLQVTTYLNLVNMSSMEFTNIYAVEHLKYSFEKCAPW 1276
 DB 1142 FAVISRHSLSAGLVGSLVSVSLQVTTYLNLVNMSSMEFTNIYAVEHLKYSFEKCAPW 1201

QY 1277 QIETAPSSMPQVGRVEFPNVCYLRVEDLDPVLRHINVTINGEKXGIYRTAGKSSL 1336
 DB 1202 QIETAPSSMPQVGRVEFPNVCYLRVEDLDPVLRHINVTINGEKXGIYRTAGKSSL 1261

QY 1337 TLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMLDPFSQYSD 1396
 DB 1262 TLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMLDPFSQYSD 1321

QY 1397 EBYWTSLELALHLDQFVSALPDKLDHBCAEGENLSVQOROLVCLARALLRKTKILVIDEA 1456
 DB 1322 EBYWTSLELALHLDQFVSALPDKLDHBCAEGENLSVQOROLVCLARALLRKTKILVIDEA 1381

QY 1457 TAAVDETDLIOSTIRTOPEDCVLTIAHRLMTIMYETVIVLDEKEIOYGAPSDLLQ 1516
 DB 1382 TAAVDETDLIOSTIRTOPEDCVLTIAHRLMTIMYETVIVLDEKEIOYGAPSDLLQ 1441

QY 1517 QRGIFYSMADAGIV 1531
 DB 1442 QRGIFYSMADAGIV 1456

RESULT 7
 Q6UR05 CANFA PRELIMINARY; PRT; 1531 AA.
 ID Q6UR05 CANFA PRELIMINARY;
 AC Q6UR05;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Multidrug resistance-associated protein 1.
 GN Name=MRP1;
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 OC NCBI_TaxID=9615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22404678; PubMed=12516967;
 RA Ma L., Prate S.B., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;
 RT "Identification and characterization of the canine multidrug
 resistance-associated protein.";
 RL Mol. Cancer Ther. 1:1335-1342(2002).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
RA Ma L., Pratt S.E., Cao J., Danczig A.H., Moore R.E., Slagac C.A.;
RL Submitted (Aug-2003) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AY63728; AAC3348.1; -, mRNA.
DR Ensembl: ENSCAFG0000018208; Canis familiaris.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR011527; ABC_membrane_1.
DR InterPro: IPR01140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR005292; MRP_assoc.
DR InterPro: IPR00719; Prot_kinase.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SMO0382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS50929; ABC_TWIF; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS0107; PROTEIN_KINASE_ATP; UNKNOWN 1.
DR ATP-binding; Nucleotide-binding; Repeat; Transport.
KW SEQUENCE 1531 AA; 171792 MW; 746361A71C6158BD CRC64;

Query Match 73.4%; Score 7272; DB 2; Length 1531;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;

QY 1 MALRGFSADSDSLMDNMNTMTNTPDPTKCFONTUVMYPCFTLACFPFFELYLSRH 60
DB 1 MALRGFSADSDSLPFWMDVSMNTSNPDFTKCFONTUVMWPCCLMCFPFELYLSRH 60
QY 61 DRGIQMTPLNKRTALGFLMIWCMADLFYSFMERSGFIAPAFYLSPTLLGITTLA 120
DB 61 DRGIQMTPLNKRTALGFLMIWCMADLFYSFMERSGFIAPAFYLSPTLLGITTLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFEDITFYVYS 180
DB 121 TFLIQLERRKGVSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFEDITFYVYS 180
QY 181 LLLIQVLVSCFSDSPLEFSETIHDNPPCESSASLSITTFWMTGLVGRYQPLESD 240
DB 181 LLLIQVLVSCFSDSPLEFSETIHDNPPCESSASLSITTFWMTGLVGRYQPLESD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPQKEMNSLRFVLYKTFGPYFLMSFFKAIHDLMMFSGPOILKLKIFVNDTKAPD 360
DB 301 IVKSPQKEMNSLRFVLYKTFGPYFLMSFFKAIHDLMMFSGPOILKLKIFVNDTKAPD 360
QY 361 WQGFYTYLLFVYTAQLQTLVHQYFHICFVSGMRKTAIVIGAVYKALVITNSARKSTV 420
DB 361 WQGFYTYLLFVYTAQLQTLVHQYFHICFVSGMRKTAIVIGAVYKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAORFMDLATYINMISAPLOVTLAYLMLNGSPVLGAVVMVLMVNVN 480
DB 421 GEIYNLMSVDAORFMDLATYINMISAPLOVTLAYLMLNGSPVLGAVVMVLMVNVN 480
QY 481 AVMAKTKTYOVANHKSKDNRIKLMNEILNGIKVLKLYAMBLAFKDYALIROBELKYLK 540
DB 481 AVMAKTKTYOVANHKSKDNRIKLMNEILNGIKVLKLYAMBLAFKDYALIROBELKYLK 540
QY 541 KSAIYLSAVGTWTWCTPPLVALCTPAVYVTTIDENNILDAQFVSLAFNLIRPLNLTLP 600
DB 541 KSAIYLSAVGTWTWCTPPLVALCTPAVYVTTIDENNILDAQFVSLAFNLIRPLNLTLP 600

QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGNTSITVNAATFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGNTSITVNAATFTWARSDDPT 660
QY 661 LMGITTSIPGALVAVVGVQCGKSSLSLLALMDKVEGHVAIKGSVAAYVPOQAWIQND 720
DB 661 LMGITTSIPGALVAVVGVQCGKSSLSLLALMDKVEGHVAIKGSVAAYVPOQAWIQND 720
QY 721 SLRENIILFGQLEPPRYRYIOACALLPDEIILPSGDRTEIGEGKVLSSGQKORVSLAR 780
DB 721 SLRENIILFGQLEPPRYRYIOACALLPDEIILPSGDRTEIGEGKVLSSGQKORVSLAR 780
QY 781 AVYSNADIVLFDPLSAVDAVGHKIPENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYCDSDIYLPDPLSAVDAVGHKIPENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGKISIMSGSYQELARDDGAPAFLLTYASTBOEQAENGVTGVSQPGKEAKOMENGM 900
DB 841 MTGKISIMSGSYQELARDDGAPAFLLTYASTBOEQAENGVTGVSQPGKEAKOMENGM 900
QY 901 LVYDSAGKOLOROLSSSSYSGDISRHNSHTAELOKAEAKKEBTWKLMEADKAOTGVKL 960
DB 901 LVYDSAGKOLOROLSSSSYSGDISRHNSHTAELOKAEAKKEBTWKLMEADKAOTGVKL 960
QY 961 SVYVDYMKALGLFISFLSIFLFCMCHVASALSNYMLSLMTDDPIVNGTOEHTKVRLSVYG 1020
DB 961 SVYVDYMKALGLFISFLSIFLFCMCHVASALSNYMLSLMTDDPIVNGTOEHTKVRLSVYG 1020
QY 1021 ALGISOGIANGVGSMAVSIIGIILASRCLHYDLHSILRSMPSPFERPSGULVNRBSKEL 1080
DB 1021 ALGISOGIANGVGSMAVSIIGIILASRCLHYDLHSILRSMPSPFERPSGULVNRBSKEL 1080
QY 1081 PTVDSMIPVYIKMFGSLFNVIACIITLAPLAIIPPLGIYFVQRFYVASSROL 1140
DB 1081 PTVDSMIPVYIKMFGSLFNVIACIITLAPLAIIPPLGIYFVQRFYVASSROL 1140
QY 1141 KRLSEVSRSPPYSHFNSTLLGVSVIRAFEEQEREITHOSDKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSEVSRSPPYSHFNSTLLGVSVIRAFEEQEREITHOSDKVDENOKAYPSIVANRWLA 1200
QY 1201 VRELCVNCIYLPALFAVISRHSLSAGLVGSYSLOVTTYINMLVRRSMSEMETNIVA 1260
DB 1201 VRELCVNCIYLPALFAVISRHSLSAGLVGSYSLOVTTYINMLVRRSMSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOENAPPSWPOVGRVFRYVCLRYRBDLPVLRHINTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOENAPPSWPOVGRVFRYVCLRYRBDLPVLRHINTINGG 1320
QY 1321 EKVGIVRTGAGKSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIILIPQDVL 1380
DB 1321 EKVGIVRTGAGKSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIILIPQDVL 1380
QY 1381 SGLSRANLDPDSQYSDDEWVTSLELAHLKOFVSLPKLDHECAEGBNLSVGOQOLVCL 1440
DB 1381 SGLSRANLDPDSQYSDDEWVTSLELAHLKOFVSLPKLDHECAEGBNLSVGOQOLVCL 1440
QY 1441 ARLALRKTLIVLEAPRAVDELTDDLIQSTIRQPDCTVLTAAHRLNTIMDTRYIVL 1500
DB 1441 ARLALRKTLIVLEAPRAVDELTDDLIQSTIRQPDCTVLTAAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEGAPSDLIQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIRREGQPSDLIQORGLFYSMAKDAGLV 1531

RESULT 8
059G19 HUMAN
ID 059G19 HUMAN PRELIMINARY; PRT; 1439 AA.
AC 059G19
DT 10-MAY-2005 (TRENBLrel. 30, Created)
DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)

DE ATP-binding cassette, sub-family C, member 1 isoform 1 variant
(Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Tsuchi Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title."
CC Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AB009120; BAD92357.1; -, mRNA.
DR Ensembl: ENSG00000103222; Homo sapiens.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:005524; F: ATP binding; IEA.
DR GO: GO:0016887; F: ATPase activity; IEA.
DR GO: GO:0042626; F: ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0006106; F: nucleotide binding; IEA.
DR GO: GO:0006810; P: transport; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR011527; ABC_membrane_1.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR005292; MRP_assoc.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRfam: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS00929; ABC_TMIF; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS00107; PROTEIN KINASE_ATP; UNKNOWN_1.
DR ATP-binding; Nucleotide-binding.
KW NON_TER
FT SEQUENCE 1 1439 AA; 160439 MW; 090D881ADBFB3077 CRC64;
SQ
Query Match 73.1%; Score 7242; DB 2; Length 1439;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1428; Conservative 0; Mismatches 1; Indels 10; Gaps 1;
QY 103 APFLVSPFLIGITTLATFLIOLERRKGVSSGIMLTFVALVCAALILRSKIMTALK 162
DB 1 APFLVSPFLIGITTLATFLIOLERRKGVSSGIMLTFVALVCAALILRSKIMTALK 60
QY 163 EDNOVLFPRITFYVYFSLLLIOLVLSCFSDRSPLPSETHDNPCESSASFLSRITFW 222
DB 61 EDNOVLFPRITFYVYFSLLLIOLVLSCFSDRSPLPSETHDNPCESSASFLSRITFW 120
QY 223 WITGLIVRGYRQPLEGSDMLSLNKEPTSEOVVPLVYKMKKECAKTRKQVKKVYSSKDP 282
DB 121 WITGLIVRGYRQPLEGSDMLSLNKEPTSEOVVPLVYKMKKECAKTRKQVKKVYSSKDP 160
QY 283 AQRKSSKVDANEVEALIVKSPQKWNPSLFLVLYKTFGPYFLMSFFPKAIDHLMFSG 342
DB 181 AQRKSSKVDANEVEALIVKSPQKWNPSLFLVLYKTFGPYFLMSFFPKAIDHLMFSG 240
QY 343 POLTKLILKFNVTTPKAPDMNGFYFYVLLPTTACLOTLLVHQQYHICVSGMRKTAIVIGA 402
DB 241 POLTKLILKFNVTTPKAPDMNGFYFYVLLPTTACLOTLLVHQQYHICVSGMRKTAIVIGA 300
QY 403 VYRKALVITNSARKSSITGEIVNIMSYDAQRFMDLATYINMISAPLOVILALYLLMLN 462
DB 301 VYRKALVITNSARKSSITGEIVNIMSYDAQRFMDLATYINMISAPLOVILALYLLMLN 360
QY 463 GPSVLAGVAVNMLVNPNAVMAMKTKTYQVAHKSKDNRKLMNEILNGIKVTKYAMEL 522
DB 361 GPSVLAGVAVNMLVNPNAVMAMKTKTYQVAHKSKDNRKLMNEILNGIKVTKYAMEL 420

QY 523 AFKDKVLAHQEELKVLKKSAYLSAVGTFPTWCPTFLVALCTFAVYVITDENNLDAQTA 582
DB 421 AFKDKVLAHQEELKVLKKSAYLSAVGTFPTWCPTFLVALCTFAVYVITDENNLDAQTA 480
QY 583 FVSLALFNILRFPNLILPWVSSIVQASVSLKRLIFLSHELEPDSIERRPVQGGGTN 642
DB 481 FVSLALFNILRFPNLILPWVSSIVQASVSLKRLIFLSHELEPDSIERRPVQGGGTN 540
QY 643 SITVRNATFTWASDDPTLNGITFSPLEGALVAVVGVGGCKSSLSALLAENDKRGHV 702
DB 541 SITVRNATFTWASDDPTLNGITFSPLEGALVAVVGVGGCKSSLSALLAENDKRGHV 600
QY 703 AIKGSVAIVPQAMQINDSLRENILFGCCOLEEYRYSVIOACALLPDLLETPSGDRFETG 762
DB 601 AIKGSVAIVPQAMQINDSLRENILFGCCOLEEYRYSVIOACALLPDLLETPSGDRFETG 660
QY 763 EKGVNLSGGQKQVSLARAVYSNADITLPDPLSAVDAAVGHKIFENVIGPKGMLKNKTR 822
DB 661 EKGVNLSGGQKQVSLARAVYSNADITLPDPLSAVDAAVGHKIFENVIGPKGMLKNKTR 720
QY 823 ILVTHSMSTYLPQVDVITVNSGGKISMGSYQELIANDGAFELRTYASTEOEDAEN- 881
DB 721 ILVTHSMSTYLPQVDVITVNSGGKISMGSYQELIANDGAFELRTYASTEOEDAEN- 780
QY 882 -----GVTVSGPKGKAKOMENGMLVTDSAGKOLQOLSSSSSGDIPRHNSTA 932
DB 781 STWDEBEAGVTVSGPKGKAKOMENGMLVTDSAGKOLQOLSSSSSGDIPRHNSTA 840
QY 933 ELQKAEAKKEETKMLEADKAQGTQVYLSYMDYMKALGLIFSLSTFLPMCHVSLAS 992
DB 841 ELQKAEAKKEETKMLEADKAQGTQVYLSYMDYMKALGLIFSLSTFLPMCHVSLAS 900
QY 993 NYWLSMTDDPIVNGTQHTKVLSTYVYALGSGISGIAVFGYSMAVSTGILLASRCLHVDL 1052
DB 901 NYWLSMTDDPIVNGTQHTKVLSTYVYALGSGIAVFGYSMAVSTGILLASRCLHVDL 960
QY 1053 LHSILSPMSFFERTSGMLVNRFSKELDVTDSMIPEVITKMFNGSLPNVIGACITVILLAT 1112
DB 961 LHSILSPMSFFERTSGMLVNRFSKELDVTDSMIPEVITKMFNGSLPNVIGACITVILLAT 1020
QY 1113 PIAAIIIPPLGLIFPYQRFYVYVSSRQLKRLLESVSRSPVYSHNETLLGVSVIRAFPEOE 1172
DB 1021 PIAAIIIPPLGLIFPYQRFYVYVSSRQLKRLLESVSRSPVYSHNETLLGVSVIRAFPEOE 1080
QY 1173 RFIHQSDLVKVDENQKAYPSIVANRWLAVERLCEVGNCTVLPALPAVISHSLSAGLVGL 1232
DB 1081 RFIHQSDLVKVDENQKAYPSIVANRWLAVERLCEVGNCTVLPALPAVISHSLSAGLVGL 1140
QY 1233 SVYSLSQVTTYLNMLVRMSSEMETNIVAVBRLEKYEETKEAWMOIOTAPPSWQVGR 1292
DB 1141 SVYSLSQVTTYLNMLVRMSSEMETNIVAVBRLEKYEETKEAWMOIOTAPPSWQVGR 1200
QY 1293 VEFPRNYCLARYEDLDVFLRHINVTINGGEKVGIVGRGTGAKSSLTJGLFPIINSAGEII 1352
DB 1201 VEFPRNYCLARYEDLDVFLRHINVTINGGEKVGIVGRGTGAKSSLTJGLFPIINSAGEII 1260
QY 1353 IDGINIAKILGHLDRFKITIIIPQDPVLFGSGLRMLNLDPSQYSDEEVMWTSLELAHKDPV 1412
DB 1261 IDGINIAKILGHLDRFKITIIIPQDPVLFGSGLRMLNLDPSQYSDEEVMWTSLELAHKDPV 1320
QY 1413 SALPDLDBHECAEGGNSLVGQROLVCLARALLRKIKIIVLDEATAVLDLETDLLQSTI 1472
DB 1321 SALPDLDBHECAEGGNSLVGQROLVCLARALLRKIKIIVLDEATAVLDLETDLLQSTI 1380
QY 1473 RTQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEGAFSPDILQOGLFFYSMAKDAGLV 1531
DB 1381 RTQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEGAFSPDILQOGLFFYSMAKDAGLV 1439
RESULT 9
Q8HXQ5_BOVIN
ID Q8HXQ5_BOVIN PRELIMINARY; PRT; 1530 AA.
AC Q8HXQ5;

DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Multidrug resistance protein 1.
GN Name=MRP1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;
RA Taguchi Y., Saeki K., Komano T.;
RT "Functional analysis of MRP1 cloned from bovine.";
RL FEBS Lett. 521:211-213(2002).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AB082124; BAC15550.1; -, mRNA.
DR HSSP: P08716; IMTO.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0006106; F:nucleotide binding; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR011527; ABC_membrane_1.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR005292; MRP_assoc.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS50929; ABC_TMIF; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1530 AA; 171666 MW; AAE4F92ED7832703 CRC64;

Query Match 73.0%; Score 7230.5; DB 2; Length 1530;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;

QY 1 MALGFCSDAGSDLLMDMNMNTMTNSNPFTKCFQNTVAVWPCFYLMACPEFYLYLSRH 60
DB 1 MALMDPFCSDVSDSLFMENNVMTNSNPDFTKCFQNTVAVWPCSTYLMVCFEYFYLYLSHH 60

QY 61 DRGYIQMTPLNKKTALGFLIMIVCWADLFYSFWEKRGIFLAPVFLVSPITLIGITLLA 120
DB 61 DRGYIQMTPLNKKTALGFLIMIVCWADLFYSFWEKRGIFLAPVFLVSPITLIGITLLA 120

QY 121 TFLIQLEBRKGVSSGIMLTWVALVCAAILNSKIMTALKEDAVDLPDITFYVES 180
DB 121 TFLIQLEBRKGVSSGIMLTWVALVCAAILNSKIMTALKEDAVDLPDITFYVES 180

QY 121 TFLIQLEBRKGVSSGIMLTWVALVCAAILNSKIMTALKEDAVDLPDITFYVES 180
DB 121 TFLIQLEBRKGVSSGIMLTWVALVCAAILNSKIMTALKEDAVDLPDITFYVES 180

QY 181 LLLIQVLVSCSDSPLEFSETIHDNPPCESASFLSITFMWITGLVRYRPLGSSD 240
DB 181 LLLIQVLVSCSDSPLEFSETIHDNPPCESASFLSITFMWITGLVRYRPLGSSD 240

QY 181 LLLIQVLVSCSDSPLEFSETIHDNPPCESASFLSITFMWITGLVRYRPLGSSD 240
DB 181 LLLIQVLVSCSDSPLEFSETIHDNPPCESASFLSITFMWITGLVRYRPLGSSD 240

QY 241 LMSLNKEDTSBQVAVPVLMKMKCECAKTRKQPVKVYSSKDPAPCKESSKVDANBEVAL 300
DB 241 LMSLNKEDTSBQVAVPVLMKMKCECAKTRKQPVKVYSSKDPAPCKESSKVDANBEVAL 300

QY 241 LMSLNKEDTSBQVAVPVLMKMKCECAKTRKQPVKVYSSKDPAPCKESSKVDANBEVAL 300
DB 241 LMSLNKEDTSBQVAVPVLMKMKCECAKTRKQPVKVYSSKDPAPCKESSKVDANBEVAL 300

QY 301 IVKSPQKEWNSLFRKVLVYKTFGPLYFMSFFKAHIDLMFSGPOLIKLIFVNDTKAPD 360
DB 301 IVKSPQKEWNSLFRKVLVYKTFGPLYFMSFFKAHIDLMFSGPOLIKLIFVNDTKAPD 360

QY 301 IVKSPQKEWNSLFRKVLVYKTFGPLYFMSFFKAHIDLMFSGPOLIKLIFVNDTKAPD 360
DB 301 IVKSPQKEWNSLFRKVLVYKTFGPLYFMSFFKAHIDLMFSGPOLIKLIFVNDTKAPD 360

QY 361 WQGFYTYVTLFVTACIQTLVJLHOYFHTCFVSGMKRKTNAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTYVTLFVTACIQTLVJLHOYFHTCFVSGMKRKTNAVIGAVYRKALVITNSARKSTV 420

QY 361 WQGFYTYVTLFVTACIQTLVJLHOYFHTCFVSGMKRKTNAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTYVTLFVTACIQTLVJLHOYFHTCFVSGMKRKTNAVIGAVYRKALVITNSARKSTV 420

QY 421 GEIVNLSVDAQRPNDLATYTNMTWSAPLOVYIALLYLMLNLGSPVLAGVAVWMLMPVN 480
DB 421 GEIVNLSVDAQRPNDLATYTNMTWSAPLOVYIALLYLMLNLGSPVLAGVAVWMLMPVN 480

QY 481 AVNAAKTKTYQVAHMSKDKRIRIKLMEILNGIRVYLKLYAMELAFKDVLAIRBELVYLK 540
DB 481 AVNAAKTKTYQVAHMSKDKRIRIKLMEILNGIRVYLKLYAMELAFKDVLAIRBELVYLK 540

QY 541 KSAVLSAVGTFTWVCPFLVALCTFAVYVYTIIDENNIIDAQTAVALFNILRPPLNLP 600
DB 541 KSAVLSAVGTFTWVCPFLVALCTFAVYVYTIIDENNIIDAQTAVALFNILRPPLNLP 600

QY 601 MVTSSIVQASVSLKRLRIFLSHBELEPDSIERRPVYKDGCGGTNSTYRNAFTWARSPPPT 660
DB 601 MVTSSIVQASVSLKRLRIFLSHBELEPDSIERRPVYKDGCGGTNSTYRNAFTWARSPPPT 660

QY 661 LMGITFSIPGALVAVVGVGCGKSSLSLSLMBEMDVEGHVAIKGSVAVYVPOQAWIQND 720
DB 661 LMGITFSIPGALVAVVGVGCGKSSLSLSLMBEMDVEGHVAIKGSVAVYVPOQAWIQND 720

QY 721 SLRENILFGQLEBPYRVSIVQACALLPDEILPSGDRTEIGEGVNLSGGQKQSVSLAR 780
DB 721 SLRENILFGQLEBPYRVSIVQACALLPDEILPSGDRTEIGEGVNLSGGQKQSVSLAR 780

QY 781 AVYSNADITYLPDPLSLAVDAHVGKHFENYIGRKMLKNTRLIYTHSMGYLPQVDYIIV 840
DB 781 AVYCDSDVYLLDPLSLAVDAHVGKHFENYIGRKMLKNTRLIYTHSMGYLPQVDYIIV 840

QY 841 MSGGKISSEMSYQELIARDGAPAFPLRTVASTBOEDAEENGVTGVSFGPGKAQKQMGNG 900
DB 841 MSGGKISSEMSYQELIARDGAPAFPLRTVASTBOEDAEENGVTGVSFGPGKAQKQMGNG 900

QY 901 LVYDSAGKQLQROLSSSSYSYSGDISRHNSSTAELOKAEKKEETWKLMEADKQTCGVKL 960
DB 901 LVYDSAGKQLQROLSSSSYSYSGDISRHNSSTAELOKAEKKEETWKLMEADKQTCGVKL 960

QY 961 SVYDWYKAIQGLFISFISIFLPMCNHVSALASNYWLSMTDDPIVNGTOHTKRLSVYVG 1020
DB 961 SVYDWYKAIQGLFISFISIFLPMCNHVSALASNYWLSMTDDPIVNGTOHTKRLSVYVG 1020

QY 1021 ALGISOGIAPFGYSMAVNSIGIILASRCLHYDLHSILRSMPSPRETPSGNLVNRFSKEL 1080
DB 1021 ALGISOGIAPFGYSMAVNSIGIILASRCLHYDLHSILRSMPSPRETPSGNLVNRFSKEL 1080

QY 1081 DTVDSMIPVYIKMFMGSLFNVIGACIYIILATPIAIIIPPLGIYFVVOFYVASSROL 1140
DB 1081 DTVDSMIPVYIKMFMGSLFNVIGACIYIILATPIAIIIPPLGIYFVVOFYVASSROL 1140

QY 1140 KRLSVSRSPVYSHFNETLLGVSVYIRAFEEQERFIROSDEKVDENQAYPSIVANRWLA 1199
DB 1140 KRLSVSRSPVYSHFNETLLGVSVYIRAFEEQERFIROSDEKVDENQAYPSIVANRWLA 1199

QY 1201 VRLFCVNCIYLPALFAVISRHSLSAGLYGLSVSYSLQVTTYVNLVRSSEMETNIVA 1260
DB 1201 VRLFCVNCIYLPALFAVISRHSLSAGLYGLSVSYSLQVTTYVNLVRSSEMETNIVA 1260

QY 1260 VERLKEVSETEKEAPMIOIETAPSSWPQYGRVFPYVYCYRREDDLPFVLRHIVTINGG 1320
DB 1260 VERLKEVSETEKEAPMIOIETAPSSWPQYGRVFPYVYCYRREDDLPFVLRHIVTINGG 1320

QY 1320 VERLKEVSETEKEAPMIOIETAPSSWPQYGRVFPYVYCYRREDDLPFVLRHIVTINGG 1320
DB 1320 VERLKEVSETEKEAPMIOIETAPSSWPQYGRVFPYVYCYRREDDLPFVLRHIVTINGG 1320

QY 1380 SSGIRLNMULDPSQYSDDEWVTSLELALKQFVSLPKXLDHECCAGGBENSVGQROLYCL 1439
DB 1380 SSGIRLNMULDPSQYSDDEWVTSLELALKQFVSLPKXLDHECCAGGBENSVGQROLYCL 1439

QY 1441 ARAILRKTKILVDEAFAVDAVLETDLLIQSTIRIQFDDCTVLTIAHRLNTIMDTRYIVL 1500
DB 1441 ARAILRKTKILVDEAFAVDAVLETDLLIQSTIRIQFDDCTVLTIAHRLNTIMDTRYIVL 1500

QY 1501 DKGEIOBYGAPSDLLQORGLFYSMAXDAGLV 1531
DB 1501 DKGEIOBYGAPSDLLQORGLFYSMAXDAGLV 1531

QY 977 LSTFLPCNHSALASYMTSLMTDDPIVNGTQHTKRLSVYCALGISOGIAVFGYSMA 1036
 Db 846 LSTFLPCNHSALASYMTSLMTDDPIVNGTQHTKRLSVYCALGISOGIAVFGYSMA 905
 QY 1037 VSIIGGLIASHRCIHDVLSHLSRSPSPFPSPCNLVNRSKELDTYDSMTPEVTKPMFG 1096
 Db 906 VSIIGGLIASHRCIHDVLSHLSRSPSPFPSPCNLVNRSKELDTYDSMTPEVTKPMFG 965
 QY 1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFVORFYVASSRQKLESYSRSPVYSHFN 1156
 Db 966 SLFNVIGACIVILLATPIAIIIPPLGLIYFVORFYVASSRQKLESYSRSPVYSHFN 1025
 QY 1157 ETLIGSVIRAFEPQERFIHOSDKVDENQAKYPSIYANMVLAVLECVANCIVLPAAL 1216
 Db 1026 ETLIGSVIRAFEPQERFIHOSDKVDENQAKYPSIYANMVLAVLECVANCIVLPAAL 1095
 QY 1217 FAVSRHSLSAGVIGSVSYLOVTTYNLNVRSSEMETNIIVVERLKXESSEKAPW 1276
 Db 1086 FAVSRHSLSAGVIGSVSYLOVTTYNLNVRSSEMETNIIVVERLKXESSEKAPW 1145
 QY 1277 QIOETAPSSWPQVGRVFRNYCLRYREDLDFVLRIHNVITNGEKVIGRTGAGKSSL 1336
 Db 1146 QIOETAPSSWPQVGRVFRNYCLRYREDLDFVLRIHNVITNGEKVIGRTGAGKSSL 1205
 QY 1337 TLGIFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPDPVPLFGSIRNMLDPFSQYSD 1396
 Db 1206 TLGIFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPDPVPLFGSIRNMLDPFSQYSD 1265
 QY 1397 EEWVTSLEIHLKDFVSLPKLHECAGEENISVQGRQVLCARALRLRTKTLVDEA 1456
 Db 1266 EEWVTSLEIHLKDFVSLPKLHECAGEENISVQGRQVLCARALRLRTKTLVDEA 1385
 QY 1457 TAAVDLETDLDLQSTIRTOFEDCTVLTAHRLNTIMDTRYVLDKGEIOEGYGAPSDLQ 1516
 Db 1326 TAAVDLETDLDLQSTIRTOFEDCTVLTAHRLNTIMDTRYVLDKGEIOEGYGAPSDLQ 1385
 QY 1517 QRGIFYSMADAGLV 1531
 Db 1386 QRGIFYSMADAGLV 1400

RESULT 11
 MRP1_MOUSE STANDARD; PRT; 1528 AA.
 AC 035379;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DE 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).
 GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mrp;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Muscle;
 RX MEDLINE=96251691; PubMed=6649356;
 RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;
 RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter.";
 RL Mol. Pharmacol. 49:962-971(1996).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STEAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=1246851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kanukawa T., Adachi J., Bono H., Kondo S., Nakai I., Osato N., Saico R., Suzuki H., Yamataka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schriber L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.M., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J., Grimmond S., Gusticich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Konegaya A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagasima T., Numata K., Okido T., Pavan W.J., Pertez G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., Savelin A., Schneider C., Sempke C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wyntlaw-Boris A., Yanagisawa M., Yang L., Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hitozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K., Shiraki T., Maki K., Kawai J., Aizawa K., Arikawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;
 RA "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 CC -1- FUNCTION: May participate directly in the active transport of drugs into subcellular organelles or influence drug distribution indirectly (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
 CC -1- SIMILARITY: Contains 2 ABC transporter domains.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC -----
 DR EMBL: AF022908; AAC80938.1; -; mRNA.
 DR EMBL: AK029876; BAC26654.1; -; mRNA.
 DR HSSP: P08716; IMT0.
 DR Ensembl: ENSMUSG0000023088; Mus musculus.
 DR MGI: MGI:102676; Abcc1.
 DR GO: GO:0005887; C: integral to plasma membrane; IDA.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR011527; ABC_membrane_1.
 DR InterPro: IPR001140; ABC_TM_transp.
 DR InterPro: IPR003439; ABC_transp_like.
 DR InterPro: IPR005292; MRP_assoc.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PD000006; ABC_transporter; 2.
 DR SMART: SM00382; AAA_2.
 DR Trifam: TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE: PS50929; ABC_TM1F; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE: PS50893; ABC_TRANSPORTER_2; 2.
 KW ATP-binding; Glycoprotein; Nucleotide-binding; Repeat; Transmembrane; Transport.
 KM TOPO_DOM 1 33 Extracellular (By similarity).
 FT 1 (By similarity).
 FT TRANSMEM 34 54 Cytoplasmic (By similarity).
 FT TOPO_DOM 55 74 Extracellular (By similarity).
 FT TRANSMEM 75 95 Extracellular (By similarity).
 FT TOPO_DOM 96 100 Extracellular (By similarity).
 FT TRANSMEM 101 121 Extracellular (By similarity).
 FT TOPO_DOM 122 133 Cytoplasmic (By similarity).
 FT TRANSMEM 134 154 Extracellular (By similarity).
 FT TOPO_DOM 155 172 Extracellular (By similarity).
 FT TRANSMEM 173 193 Extracellular (By similarity).
 FT TOPO_DOM 194 317 Cytoplasmic (By similarity).
 FT TRANSMEM 318 338 Extracellular (By similarity).
 FT TOPO_DOM 339 364 Extracellular (By similarity).

FT	TRANSMEM	365	385	7 (By similarity).
FT	TOPO_DOM	386	441	Cytoplasmic (By similarity).
FT	TRANSMEM	442	462	8 (By similarity).
FT	TOPO_DOM	463	465	Extracellular (By similarity).
FT	TRANSMEM	466	486	9 (By similarity).
FT	TOPO_DOM	487	548	Cytoplasmic (By similarity).
FT	TRANSMEM	549	569	10 (By similarity).
FT	TOPO_DOM	590	591	Extracellular (By similarity).
FT	TRANSMEM	592	612	11 (By similarity).
FT	TOPO_DOM	613	663	Cytoplasmic (By similarity).
FT	TRANSMEM	664	984	12 (By similarity).
FT	TOPO_DOM	985	1022	Extracellular (By similarity).
FT	TRANSMEM	1023	1043	13 (By similarity).
FT	TOPO_DOM	1044	1086	Cytoplasmic (By similarity).
FT	TRANSMEM	1087	1107	14 (By similarity).
FT	TOPO_DOM	1108	1108	Extracellular (By similarity).
FT	TRANSMEM	1109	1129	15 (By similarity).
FT	TOPO_DOM	1130	1200	Cytoplasmic (By similarity).
FT	TRANSMEM	1201	1221	16 (By similarity).
FT	TOPO_DOM	1222	1223	Extracellular (By similarity).
FT	TRANSMEM	1224	1244	17 (By similarity).
FT	TOPO_DOM	1245	1528	Cytoplasmic (By similarity).
FT	DOMAIN	326	609	ABC transmembrane type-1.
FT	DOMAIN	644	868	ABC transporter 1.
FT	DOMAIN	971	1253	ABC transmembrane type-1.
FT	DOMAIN	1290	1524	ABC transporter 2.
FT	NP_BIND	678	685	ATP 1 (Potential).
FT	NP_BIND	1324	1331	ATP 2 (Potential).
FT	CARBOHYD	19	19	N-linked (GLCNAC. . .) (Potential).
FT	CARBOHYD	1003	1003	N-linked (GLCNAC. . .) (Potential).
SEQ	SEQUENCE	1528 AA;	171185 MW;	68FD13667D61DBB CRC64;

Query Match 70.7%; Score 7002.5; DB 1; Length 1528; Best Local Similarity 88.0%; Pred. No. 0; Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY	1	MALRGFCSADGSDPLMDMNTMTNPDFTKCFONTLVVWPCYTLACFPFYLYSRH	60
DB	1	MALRFSFCSADGSDPLMDMNTMTNPDFTKCFONTLVVWPCYTLACFPFYLYSRH	60
QY	61	DRGYIQMTPLNKTKTALGFLMTIVCMADLFYSFWERSRGIFLAPVFLVPTLLGITTLA	120
DB	61	DRGYIQMTPLNKTKTALGFLMTIVCMADLFYSFWERSRGIFLAPVFLVPTLLGITTLA	120
QY	121	TEFLQERRRQVSSGIMLTFMWALVCAALISKTMALKEDAOVDLFRDITFYFYS	180
DB	121	TEFLQERRRQVSSGIMLTFMWALVCAALISKTMALKEDAOVDLFRDITFYFYS	180
QY	181	LTLILQVLSCFSDRSPLESTIHDPNCPSSASFLRITFMWTGLIVGYOQLSGSD	240
DB	181	LTLILQVLSCFSDRSPLESTIHDPNCPSSASFLRITFMWTGLIVGYOQLSGSD	240
QY	241	LWLNKEDTSEOVVPLVKKWKECAKTRKQPVKVS-SKDPAPRESSKVDANEEVEA	299
DB	241	LWLNKEDTSEOVVPLVKKWKECAKTRKQPVKVS-SKDPAPRESSKVDANEEVEA	299
QY	300	LIVKSPKEDREPSEFLKLYTTFGPFYLMSEFLYKALHDLMPFAGPKIIEINFNDRAP	360
DB	300	LIVKSPKEDREPSEFLKLYTTFGPFYLMSEFLYKALHDLMPFAGPKIIEINFNDRAP	360
QY	360	DMOGYFTVLLFYTAGCQTLVHOYFHI CFVSGKRITAVTIGAVYRRLVITNSARSSST	419
DB	360	DMOGYFTVLLFYTAGCQTLVHOYFHI CFVSGKRITAVTIGAVYRRLVITNSARSSST	419
QY	420	VGEIVNLMSYDAQSFMDLATYINMWSAPLOVILATLMLNLGSPVLAGAVAVMLAVPV	479
DB	420	VGEIVNLMSYDAQSFMDLATYINMWSAPLOVILATLMLNLGSPVLAGAVAVMLAVPV	479
QY	480	NATMAKTKTYQVAHMSKDNRIKLANEINGRIYVLTLYAMELAFKDKVLAIRQELKVL	539
DB	480	NATMAKTKTYQVAHMSKDNRIKLANEINGRIYVLTLYAMELAFKDKVLAIRQELKVL	539
QY	540	KKSAYLSAVGTFTWVCTPFLVALCTFAVYVTIDENNILDAQTLAFVSLATFNILRFLNLT	599
DB	540	KKSAYLSAVGTFTWVCTPFLVALCTFAVYVTIDENNILDAQTLAFVSLATFNILRFLNLT	599

DB	541	KKSAYLSAVGTFTWVCTPFLVALCTFAVYVTIDENNILDAQTLAFVSLATFNILRFLNLT	600
QY	600	PMYISSVQASVSLKRLRILSLHELEPDSIERRPVQDGGTINSITRNATFTYARSDPP	659
DB	601	PMYISSVQASVSLKRLRILSLHELEPDSIERRSISGSG--NSITVKNATFTYARSDPP	659
QY	660	TLNGITPSIEGALVAVVGVGGKSLSLALAEOMKVGHVAIKGSVAVVPOQAION	719
DB	660	TLNGITPSIEGALVAVVGVGGKSLSLALAEOMKVGHVAIKGSVAVVPOQAION	719
QY	720	DSLRENIIFGCOLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSGGQKQVSLA	779
DB	720	DSLRENIIFGCOLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSGGQKQVSLA	779
QY	780	RAYSNSDIYLPDDPLSAVDAAVGHKHI FENVIPKMKLKNKTRILVTHSMYLPQVDVII	839
DB	780	RAYSNSDIYLPDDPLSAVDAAVGHKHI FENVIPKMKLKNKTRILVTHSMYLPQVDVII	839
QY	840	VMSGKISSENGSYOELLARGAFAELFTASTEOEDAEENGVTGVSGRKAKOMENG	899
DB	840	VMSGKISSENGSYOELLARGAFAELFTASTEOEDAEENGVTGVSGRKAKOMENG	899
QY	900	MLVDSAGKQLOKOLSSSSSYSGDISRHNHSTABLOKAEKKEETMKLMAADRAQTOGVK	959
DB	900	MLVDSAGKQLOKOLSSSSSYSGDISRHNHSTABLOKAEKKEETMKLMAADRAQTOGVK	959
QY	960	LSVYMDYMKALIGLIFSLIFELFMCHVSAALASNYWLSMTWD--PIVNGTOEHTKRLSV	1018
DB	960	LSVYMDYMKALIGLIFSLIFELFMCHVSAALASNYWLSMTWD--PIVNGTOEHTKRLSV	1018
QY	1019	YGALGISQGIYAVGYSAVSIIGLILASRCHLVLLHSILASPMSPFERTSGNULVRFK	1078
DB	1019	YGALGISQGIYAVGYSAVSIIGLILASRCHLVLLHSILASPMSPFERTSGNULVRFK	1078
QY	1079	ELDTVDSMPEVVKMFNGSLFENYIACIYVLLATPAAIIIPPLGLYFPYQRYVYASSR	1138
DB	1079	ELDTVDSMPEVVKMFNGSLFENYIACIYVLLATPAAIIIPPLGLYFPYQRYVYASSR	1138
QY	1139	QLKRLSVSRSPYVSHENETLLGVSVYIARFEOERFIHQSDLVKVDENOKAYPSIVANRW	1198
DB	1139	QLKRLSVSRSPYVSHENETLLGVSVYIARFEOERFIHQSDLVKVDENOKAYPSIVANRW	1198
QY	1199	LAVRLCEVNCIYLPALPAVISRHSLSAGLVGLSYSLQVTTYANLVYRMSSEMTNI	1258
DB	1199	LAVRLCEVNCIYLPALPAVISRHSLSAGLVGLSYSLQVTTYANLVYRMSSEMTNI	1258
QY	1259	VAVERLKEVSETEKEAPMWOIQTAPPSWQVGRVFRNCLRYREDLDFVLRHINTIN	1318
DB	1259	VAVERLKEVSETEKEAPMWOIQTAPPSWQVGRVFRNCLRYREDLDFVLRHINTIN	1318
QY	1319	GGEKVGIVRTGAKSLLTGLFRINESAGEIILIDGINIAKTLGLDLRFRKTIIPDPV	1378
DB	1319	GGEKVGIVRTGAKSLLTGLFRINESAGEIILIDGINIAKTLGLDLRFRKTIIPDPV	1378
QY	1379	LFGSGLRMLNDPPSQISDEEVTSLBLAHKDFVSALPKLDBECABGGENLSVGOQOLV	1438
DB	1379	LFGSGLRMLNDPPSQISDEEVTSLBLAHKDFVSALPKLDBECABGGENLSVGOQOLV	1438
QY	1439	CLARALLRKTILVLEATVAVDLETDLLOSTIRFOPEBCYTLTAHRLNTIMDTYRI	1498
DB	1439	CLARALLRKTILVLEATVAVDLETDLLOSTIRFOPEBCYTLTAHRLNTIMDTYRI	1498
QY	1499	VLDKGEIOEYGAPEDDLQOGRGLFYSAKDAGLV	1531
DB	1499	VLDKGEIOEYGAPEDDLQOGRGLFYSAKDAGLV	1531

RESULT 12
0810E4 RAT PRELIMINARY; PRT; 1532 AA.
ID 0810E4 RAT
AC 0810E4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE ATP-binding cassette protein C1.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=SpLeen;
RA Yabuchi H., Takayanagi S., Ishikawa T.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -; SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF487549; AAC85437.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR Ensemble; ENSRNOG00000032748; Rattus norvegicus.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0006106; F:nucleotide binding; IEA.
DR GO; GO:0006101; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_TM_transp.
DR InterPro; IPR001395; Aldol_ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR Prodom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRfams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS0929; ABC_TWIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDO_KETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1532 AA; 171493 MW; 2E639F63F5A3F68 CRC64;

Query Match 70.0%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;

QY 1 MALRGFCSADGSDPLMDMNTWNTSNDFTKCFQNTVLVWPCFYLAACEFFLYLSRH 60
DB 1 MALSSFCSSDSDPLMDMNTWNTSNDFTKCFQNTVLVWPCFYLSRL 60
QY 61 DRGYIQMTPLNKTKTALGFILMIVCMADLFYSFWERSRGIFLAIVFLVSPILLGITLLA 120
DB 61 DRGYIQMTPLNKTKTALGFILMIVCMADLFYSFWERSRGIFLAIVFLVSPILLGITLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMVALVCALALIRSKIMTALKEDAQVDLFFDITFYVFS 180
DB 121 TFLIOLERRKGVSSGIMLTFMVALVCALALIRSKIMTALKEDAQVDLFFDITFYVFS 180
QY 181 LLLIQLVLSCPSDRSPLFSETIHDNPNCPRESSASFLSRITTWITGLIVRGYROPLEGSD 240
DB 181 LLLIQLVLSCPSDRSPLFSETIHDNPNCPRESSASFLSRITTWITGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSQVQVPLVYKMKKCECAKTRKQPVVYVS-SKDPNQPKSSKVDANEVEA 299
DB 241 LMSLNKEDTSQVQVPLVYKMKKCECAKTRKQPVVYVS-SKDPNQPKSSKVDANEVEA 299
QY 300 LIVESPOKEMNPSLFKYLKTFGEFYFLMSFFKAIHIDLMFSGPOLILKLFVNDYAP 359
DB 301 LIVESPOKEMNPSLFKYLKTFGEFYFLMSFFKAIHIDLMFSGPOLILKLFVNDYAP 359
QY 360 DMGQYFTVLVFTACIQTVLVHOYFHIQVSGMKIKTAIVIGAYRKALVITNSARKST 419
DB 360 DMGQYFTVLVFTACIQTVLVHOYFHIQVSGMKIKTAIVIGAYRKALVITNSARKST 419

DB 361 DMGQYFTALLFPASCIQTALVHOYFHIQVSGMKIKTAIVIGAYRKALVITNSARKST 420
QY 420 VGEIVINMSVDAQRFMDLATYINNMWSAPLOVILATLWNLNPSYLAGAVANVLWVPV 479
DB 421 VGEIVINMSVDAQRFMDLATYINNMWSAPLOVILATLWNLNPSYLAGAVANVLWVPV 480
QY 480 NAWAMKTKTYQVAHMSKDNRIKANEIINGIKVLKLYAMELAFOKQVLAIROEELKVL 539
DB 481 NAWAMKTKTYQVAHMSKDNRIKANEIINGIKVLKLYAMELAFOKQVLAIROEELKVL 540
QY 540 KKSATVSAVGTFPMVCPPLVALCTPAVYVTIENNLLDQOTAFVSLAFNUIRFPNITL 599
DB 541 KKSATVLAAGVFTWCTPFLVALSTPAVYVBEKNLLDKAKKAFSLAFNUIRFPNITL 600
QY 600 PMVSSIVQASVSLKRIIFLSHELEPDSIERPVDGCGTNSITVRNAFTWARSDDP 659
DB 601 PMVSSIVQASVSLKRIIFLSHELEPDSIERPVDGCGTNSITVRNAFTWARSDDP 660
QY 660 TLNGITSIEGALVAVGVGCGKSLSLALLAEMDKVEGHVAIKQSVAVVPOQAMION 719
DB 661 TLNGITPAIDGALVAVGVGCGKSLSLALLAEMDKVEGHVTLKQSVAVVPOQAMION 720
QY 720 DSIJRENTLPGQLEEPYRSVIACALLPLEITLPSGDRTEIGRGVNLGGQKQVSIA 779
DB 721 DSIJRENTLPGQLEEPYRSVIACALLPLEITLPSGDRTEIGRGVNLGGQKQVSIA 780
QY 780 RAVYSNADIVLFDPLSAVDAAHVKHIFENVIGPKGMKNKTRILVTHSMASYLPQVDVIL 839
DB 781 RAVYSNADIVLFDPLSAVDAAHVKHIFENVIGPKGMKNKTRILVTHSMASYLPQVDVIL 840
QY 840 VMSGKRISEKGSYOELLARDGAPELRTYASTFEQEDAEENGVTGVSQPKKAKQENG 899
DB 841 VMSGKRISEKGSYOELLARDGAPELRTYASTFEQEDAEENGVTGVSQPKKAKQENG 900
QY 900 MLVYDSAGKQORLSSSSSGDISRHNSSTLEQAEAKKEBTKLMEADKQOTQOVK 959
DB 901 ILVYDAVGKPLQRLSSSSSVTNOQSSVTELQKS-GYKEBTKLMEADKQOTQOVK 959
QY 960 LSVYWDPMKAIIGLFIETSLIFLFWCNVSAALASNYMLSTWTD-PIVNGTOEHTKVALSV 1018
DB 960 LSVYWDPMKAIIGLFIETSLIFLFWCNVSAALASNYMLSTWTD-PIVNGTOEHTKVALSV 1019
QY 1019 YGALGISQGIAPVGVSAVSIIGILASRCLHVDLHLSILRSFMSFFERTSGNIVNPSK 1078
DB 1020 YGALGISQGIAPVGVSAVSIIGILASRCLHVDLHLSILRSFMSFFERTSGNIVNPSK 1079
QY 1079 ELDTVDSDMIPQVTKMFNGSLFNYIGACTIVILLATPIAIIIPPLGLYFFVQRYVASSR 1138
DB 1080 ELDTVDSDMIPQVTKMFNGSLFNYIGACTIVILLATPIAIIIPPLGLYFFVQRYVASSR 1139
QY 1139 QLRKLESVSRSPVYSHENETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRW 1198
DB 1140 QLRKLESVSRSPVYSHENETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRW 1199
QY 1199 LAVALRECVGNICVILFAALFAVIRSHSLASGLVGLSVSYLQVTTYLNMVLRMSSEMETNI 1258
DB 1200 LAVALRECVGNICVILFAALFAVIRSHSLASGLVGLSVSYLQVTTYLNMVLRMSSEMETNI 1259
QY 1259 VAVERLKEYSETKEAWQOETRAPBSMPQGVVERFRANCLARREDLPVLRHINVTIN 1318
DB 1260 VAVERLKEYSETKEAWQOETRAPBSMPQGVVERFRANCLARREDLPVLRHINVTIN 1319
QY 1319 GGEKVGIVRGAGKSLTGLPINSAGEEIIIDGINIAKGLHDLRFKTIITIPDDPV 1378
DB 1320 GGEKVGIVRGAGKSLTGLPINSAGEEIIIDGINIAKGLHDLRFKTIITIPDDPV 1379
QY 1379 LFGSGLRMMNLDPFSQYSDERWVTSLELAHLKDFFVSALPDKLDHECAEGENTSVGQROLY 1438
DB 1380 LFGSGLRMMNLDPFSQYSDERWVTSLELAHLKDFFVSALPDKLDHECAEGENTSVGQROLY 1439
QY 1439 CLARALLRKTKIILVDAATAVADLETDDLQSTIRTOFEDCTVITIAHRLNTIMDYRVI 1498
DB 1440 CLARALLRKTKIILVDAATAVADLETDDLQSTIRTOFEDCTVITIAHRLNTIMDYRVI 1499

QY 1319 GGEKVGIVGRTGACKSSITLGLFRINSSAGEEIIIDGINIAKIGLHDIRKITTIIIPDDV 1378
DB 1320 GGEKVGIVGRTGACKSSITLGLFRINSSAGEEIIIDGINIAKIGLHDIRKITTIIIPDDV 1379
QY 1379 LFPSSGLMWNLDPEGOYSDPEFWTSLSLAHLKDPVSALPDKLDHRCAGEGMLSGQRLV 1438
DB 1380 LFPSSGLMWNLDPEGOYSDPEFWTSLSLAHLKDPVSALPDKLDHRCAGEGMLSGQRLV 1439
QY 1439 CLARALRKTKILVLDENATAVLDLTDLIQSTIRTPQEDCTVLTIARLNTIMDYTRVI 1498
DB 1440 CLARALRKTKILVLDENATAVLDLTDLIQSTIRTPQEDCTVLTIARLNTIMDYTRVI 1499
QY 1499 VLDKGEIOEGYAPSDLLQQRGLFYSMAKAGLV 1531
DB 1500 VLDKGEIOEGYAPSDLLQQRGLFYSMAKAGLV 1532
RESULT 14
Q810G9_RAT PRELIMINARY; PRT; 1523 AA.
ID Q810G9_RAT PRELIMINARY; PRT; 1523 AA.
AC Q810G9;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ATP-binding cassette protein C1 variant A.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Takayanagi S., Ichikawa T.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY174892; AAC44983.1; -, mRNA.
DR HSSP; P08716; IMT0.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS0929; ABC_TM1F; 2; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1523 AA; 170505 MW; E40337051A1CB9C6 CRC64;
Query Match 69.6%; Score 6892; DB 2; Length 1523;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;
QY 1 MALGFCAGDSDPPLMDNMVNTSNPPTKCFONTLVVWVPCFYLMACFPYFLYLSRH 60
DB 1 MALGFCAGDSDPPLMDNMVNTSNPPTKCFONTLVVWVPCFYLMACFPYFLYLSRH 60
QY 61 DRGYQMTPLNKTALGFLMLIVCMADLFYSFWERSRGIFLAPVFLVSPILLGITTLA 120

DB 61 DRGYQMTPLNKTALGFLMLIVCMADLFYSFWERSRGIFLAPVFLVSPILLGITTLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMWALVCAALILRSKMTALMKEDAVDLFPEDITFYVYS 180
DB 121 TFLIOLERRKGVSSGIMLTFMWALVCAALILRSKMTALMKEDAVDLFPEDITFYVYS 180
QY 121 TFLIOLERRKGVSSGIMLTFMWALVCAALILRSKMTALMKEDAVDLFPEDITFYVYS 180
DB 121 TFLIOLERRKGVSSGIMLTFMWALVCAALILRSKMTALMKEDAVDLFPEDITFYVYS 180
QY 181 LLLIQLVLSGFSRDRPLFSETIHDNPNCPSSSASFLSRTIFWMTGLIVRGQRLPBGSD 240
DB 181 LLLIQLVLSGFSRDRPLFSETIHDNPNCPSSSASFLSRTIFWMTGLIVRGQRLPBGSD 240
QY 181 LLLIQLVLSGFSRDRPLFSETIHDNPNCPSSSASFLSRTIFWMTGLIVRGQRLPBGSD 240
DB 181 LLLIQLVLSGFSRDRPLFSETIHDNPNCPSSSASFLSRTIFWMTGLIVRGQRLPBGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKEGACATRKQPKVYVS-SKDPAPQPKSSAVDANEVEA 299
DB 241 LMSLNKEDTSEQVVPVLVKNWKEGACATRKQPKVYVS-SKDPAPQPKSSAVDANEVEA 299
QY 241 LMSLNKEDTSEQVVPVLVKNWKEGACATRKQPKVYVS-SKDPAPQPKSSAVDANEVEA 299
DB 241 LMSLNKEDTSEQVVPVLVKNWKEGACATRKQPKVYVS-SKDPAPQPKSSAVDANEVEA 299
QY 300 LIVESPOKEMNPSLFKVLYKTFPGFYPLMSFFFKAIHDLMSPGQILKLKLFVNDTKAP 359
DB 300 LIVESPOKEMNPSLFKVLYKTFPGFYPLMSFFFKAIHDLMSPGQILKLKLFVNDTKAP 359
QY 301 LIVKSHKDRDPSLFKVLKTFPGFYPLMSFFFKAIHDLMSPGQILKLKLFVNDTKAP 360
DB 301 LIVKSHKDRDPSLFKVLKTFPGFYPLMSFFFKAIHDLMSPGQILKLKLFVNDTKAP 360
QY 360 DMQGYFYTVLLFVYACLOTIVLHQYFHCIVSGNRKRTAVIAGAVYRKALVITNSARKST 419
DB 360 DMQGYFYTVLLFVYACLOTIVLHQYFHCIVSGNRKRTAVIAGAVYRKALVITNSARKST 419
QY 361 DMQGYLYTALLFVSACTIOTALHQYFHCIVTGMRIKTA VVGA VYRKALVITNSARKST 420
DB 361 DMQGYLYTALLFVSACTIOTALHQYFHCIVTGMRIKTA VVGA VYRKALVITNSARKST 420
QY 420 VGEIVNLSVDAQRFMDLATYINNIWNSAPLOVILALVLMNLGSPVLAGVAVWLVVPV 479
DB 420 VGEIVNLSVDAQRFMDLATYINNIWNSAPLOVILALVLMNLGSPVLAGVAVWLVVPV 479
QY 421 VGEIVNLSVDAQRFMDLATYINNIWNSAPLOVILALVLMNLGSPVLAGVAVWLVVPV 480
DB 421 VGEIVNLSVDAQRFMDLATYINNIWNSAPLOVILALVLMNLGSPVLAGVAVWLVVPV 480
QY 480 NAWNAMTKTYQVAMHMSKDNRIKLMEILNGIKVLYLAMELAFKQVLAIRQELKVL 539
DB 480 NAWNAMTKTYQVAMHMSKDNRIKLMEILNGIKVLYLAMELAFKQVLAIRQELKVL 539
QY 481 NAWNAMTKTYQVAMHMSKDNRIKLMEILNGIKVLYLAMELAFKQVLAIRQELKVL 540
DB 481 NAWNAMTKTYQVAMHMSKDNRIKLMEILNGIKVLYLAMELAFKQVLAIRQELKVL 540
QY 540 KKSAYLSAVGTFVWCTPPLVALCTPAVYVITIDENILDAQTAVSLAFNLIRPMLNL 599
DB 540 KKSAYLSAVGTFVWCTPPLVALCTPAVYVITIDENILDAQTAVSLAFNLIRPMLNL 599
QY 541 KKSAYLSAVGTFVWCTPPLVALCTPAVYVITIDENILDAQTAVSLAFNLIRPMLNL 600
DB 541 KKSAYLSAVGTFVWCTPPLVALCTPAVYVITIDENILDAQTAVSLAFNLIRPMLNL 600
QY 600 PMVTSIYQASVSLKRLRIFLSHEELRPDSIERBPVYDGGGTSITTRNATFTARSDPP 659
DB 600 PMVTSIYQASVSLKRLRIFLSHEELRPDSIERBPVYDGGGTSITTRNATFTARSDPP 659
QY 601 PMVTSIYQASVSLKRLRIFLSHEELRPDSIERBPVYDGGGTSITTRNATFTARSDPP 660
DB 601 PMVTSIYQASVSLKRLRIFLSHEELRPDSIERBPVYDGGGTSITTRNATFTARSDPP 660
QY 660 TLNGITTSIPGALVAVVAGVCGCKSSLSLALMEMDKVEGVAKSVAVVPOAMION 719
DB 660 TLNGITTSIPGALVAVVAGVCGCKSSLSLALMEMDKVEGVAKSVAVVPOAMION 719
QY 661 TLNGITTSIPGALVAVVAGVCGCKSSLSLALMEMDKVEGVAKSVAVVPOAMION 720
DB 661 TLNGITTSIPGALVAVVAGVCGCKSSLSLALMEMDKVEGVAKSVAVVPOAMION 720
QY 720 DLSRENILFGCOLPEPYRYVQACALLPDLIELPSGDREIIGKGVNLSGGQKORVSLA 779
DB 720 DLSRENILFGCOLPEPYRYVQACALLPDLIELPSGDREIIGKGVNLSGGQKORVSLA 779
QY 721 DLSRENILFGCOLPEPYRYVQACALLPDLIELPSGDREIIGKGVNLSGGQKORVSLA 780
DB 721 DLSRENILFGCOLPEPYRYVQACALLPDLIELPSGDREIIGKGVNLSGGQKORVSLA 780
QY 780 RAVYSNADITYLPDPLSAVDHVGKHFENVYIGKGMKAKTRILVTHSMYLPQVDYI 839
DB 780 RAVYSNADITYLPDPLSAVDHVGKHFENVYIGKGMKAKTRILVTHSMYLPQVDYI 839
QY 781 RAVYCNSDIYLLDPLSAVDHVGKHFENVYIGKGMKAKTRILVTHSMYLPQVDYI 840
DB 781 RAVYCNSDIYLLDPLSAVDHVGKHFENVYIGKGMKAKTRILVTHSMYLPQVDYI 840
QY 840 VMSGKTSKEMGSYQELARDOAFAPLRTVASTBOEDAEENGTGVYSGPKAKOMENG 899
DB 840 VMSGKTSKEMGSYQELARDOAFAPLRTVASTBOEDAEENGTGVYSGPKAKOMENG 899
QY 841 VMSGKTSKEMGSYQELARDOAFAPLRTVASTBOEDAEENGTGVYSGPKAKOMENG 900
DB 841 VMSGKTSKEMGSYQELARDOAFAPLRTVASTBOEDAEENGTGVYSGPKAKOMENG 900
QY 900 MLVYDSAGKQORQLSSSSYSYSDISRHNSTAEIQRKAEKERTKXLMKADKAQYGVK 959
DB 900 MLVYDSAGKQORQLSSSSYSYSDISRHNSTAEIQRKAEKERTKXLMKADKAQYGVK 959
QY 901 ILVYDAVAKPL-----HSVYTNQHSSTAEIQGS-GVKEFTWKLMEADKXQYGVK 950
DB 901 ILVYDAVAKPL-----HSVYTNQHSSTAEIQGS-GVKEFTWKLMEADKXQYGVK 950
QY 960 LSVYWDYKAIQLFISPLSTFLPMONHVSALASNTWLSLWTD-PIYNGTOEHYKRLSV 1018
DB 960 LSVYWDYKAIQLFISPLSTFLPMONHVSALASNTWLSLWTD-PIYNGTOEHYKRLSV 1018
QY 951 LSVYWNWYKAIQLFISPLSTFLPMONHVSALASNTWLSLWTD-PIYNGTOEHYKRLSV 1010
DB 951 LSVYWNWYKAIQLFISPLSTFLPMONHVSALASNTWLSLWTD-PIYNGTOEHYKRLSV 1010
QY 1019 YGALGIGSGLIAPVGSMAVSTIGLILASHVDLILSLRSPMSFFERTPSGNTVNRPSK 1078
DB 1019 YGALGIGSGLIAPVGSMAVSTIGLILASHVDLILSLRSPMSFFERTPSGNTVNRPSK 1078
QY 1011 YGALGIGSGLIAPVGSMAVSTIGLILASHVDLILSLRSPMSFFERTPSGNTVNRPSK 1070
DB 1011 YGALGIGSGLIAPVGSMAVSTIGLILASHVDLILSLRSPMSFFERTPSGNTVNRPSK 1070
QY 1079 ELDPVDSMIPVIVIKFMGSLFENVIGACVILATPIAIIIPPLGLIYFFQRFVASSR 1138
DB 1079 ELDPVDSMIPVIVIKFMGSLFENVIGACVILATPIAIIIPPLGLIYFFQRFVASSR 1138
QY 1071 ELDPVDSMIPVIVIKFMGSLFENVIGACVILATPIAIIIPPLGLIYFFQRFVASSR 1130
DB 1071 ELDPVDSMIPVIVIKFMGSLFENVIGACVILATPIAIIIPPLGLIYFFQRFVASSR 1130
QY 1139 QLRLESVSRSPVYSHFNETLLGVSIVIRAFEEORFTHQSGLKVDENOKAYYPSIVANRW 1198
DB 1139 QLRLESVSRSPVYSHFNETLLGVSIVIRAFEEORFTHQSGLKVDENOKAYYPSIVANRW 1198

Db 1131 QKRLSESVSRPVSHFNETLLGVSVIRAFEEORERFIROSDLKVDENQKAYPSIVANRW 1190
 Qy 1199 LAVALTECVGNCIYV.FAALFAVISRHSLSAGLVGLSVSYSLQVTTYIMLVMSSEMENI 1258
 Db 1191 LAVALTECVGNCIYV.FAALFAVISRHSLSAGLVGLSVSYSLQVTTYIMLVMSSEMENI 1250
 Qy 1259 VAVERLKEVSETEKEAPWQIOETAPPSWPOGVHVEFPNNYCLARYREDLDFLRHINVTIN 1318
 Db 1251 VAVERLKEVSETEKEASWQIOETAPPSWPHSGRVERPFDYCLRRREDLVLKINVTIE 1310
 Qy 1319 GGEKVGVIGRTGAKSSITLGLFRINSEAGEEIIIDGINAKIGLHDLRFKITTIIPODPV 1378
 Db 1311 GGEKVGVIGRTGAKSSITLGLFRINSEAGEEIIIDGINAKIGLHDLRFKITTIIPODPV 1370
 Qy 1379 LFGSLTLMNLDPEFOYSDPEEWVTSLELAHLKDPFSAIPDKLDHFCAGEGENLSTVGORQV 1438
 Db 1371 LFGSLTLMNLDPEFOYSDPEEWVTSLELAHLKDPFSAIPDKLDHFCAGEGENLSTVGORQV 1430
 Qy 1439 CLAAALRKTKILVLDENATAVDETDLIOSTIRTOPEDCVTLTIAHRLNTIMDYTRVI 1498
 Db 1431 CLAAALRKTKILVLDENATAVDETDLIOSTIRTOPEDCVTLTIAHRLNTIMDYTRVI 1490
 Qy 1499 VLDKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531
 Db 1491 VLDKGEIOEYGAPSDLLQORGLFYSMADAGLV 1523

RESULT 15
 Q5F364_CHICK
 ID Q5F364_CHICK PRELIMINARY; PRT; 1525 AA.
 AC Q5F364;
 DT 10-MAY-2005 (Tremblrel. 30, Created)
 DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
 DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=RCJMB04.32d20;
 OS Gallus gallus (Chicken).
 OC Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CH; TISSUE=Bursa;
 RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
 RA Pledy J., Carnuci P., Blasodatski A., Kostovska D., Kotler M.,
 RA "Full-length cDNAs from chicken bursa lymphocytes to facilitate
 RT genefunction analysis."
 RL Genome Biol. 6:R6(2005).
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; AJ851786; CAH65420.1; -; mRNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016887; F:ATPase activity; IEA.
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR011527; ABC_membrane_1.
 DR InterPro; IPR001140; ABC_TM_transp_1.
 DR InterPro; IPR003439; ABC_TM_transp_1.
 DR InterPro; IPR005292; MRP_assoc.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR Prodom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE; PSS0929; ABC_TMIF; 2.
 DR PROSITE; PSS0021; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PSS0893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PSS0107; PROTEIN_KINASE_ATP; UNKNOWN_1.

KW ATP-binding; Hypothetical protein; Nucleotide-binding; Repeat;
 KW Transport. 1525 AA; 170972 MW; 387246711B878FD CR664;
 SQ SEQUENCE
 Query Match 63.6%; Score 6301; DB 2; Length 1525;
 Best local similarity 77.6%; Pred. No. 0;
 Matches 1193; Conservative 176; Mismatches 148; Indels 20; Gaps 7;
 Qy 1 MALRGFSSAGSDPLMNMVNTNNTSNDPFCRCQNTLVWVPCYTLACPPFYLYSRH 60
 Db 1 MGIESLSAASAEFFWMNLTWHTENPDFQCFQNTLVWVPCYTLWCPPEAYLYLRSH 60
 Qy 61 DRGYIQMTPLNKTALGFLMTIYCMADLFYSPFERSRGIFLAPVFLVSPILGITLLA 120
 Db 61 DRGYIQMTPLNKTALGFLMTIYCMADLFYSPFERSRGIFLAPVFLVSPILGITLLA 120
 Qy 121 TFLIQERRKGVSSGIMLTFWVALCALALRSKIMTALKEDAQVDFPRDIFYVYFS 180
 Db 121 TFLIQERRKGVSSGIMLTFWVALCALALRSKIMTALKEDAQVDFPRDIFYVYFS 180
 Qy 181 LLLIQVLVSCPSDRSLFSETIHDPNPCSSSASFLSRITFWMTGLIVNGYRPLEGSD 240
 Db 181 LLLIQVLVSCPSDRSLFSETIHDPNPCSSSASFLSRITFWMTGLIVNGYRPLEGSD 240
 Qy 181 LLLVQLILSCFPEKPLFSEAVVNDPKCPFSASFLSRITFWMTGLIMIQHRRPLAKD 240
 Db 181 LLLVQLILSCFPEKPLFSEAVVNDPKCPFSASFLSRITFWMTGLIMIQHRRPLAKD 240
 Qy 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPVKVVYSSKDPAPQKESKYVDN----EE 296
 Db 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPVKVVYSSKDPAPQKESKYVDN----EE 296
 Qy 297 VEALIVKSPQKWNPSLFKVLVYKTFGPFLMSFFPKAHDLMFSGPOLIKLILKPVNDT 356
 Db 297 VEALIVKSPQKWNPSLFKVLVYKTFGPFLMSFFPKAHDLMFSGPOLIKLILKPVNDT 356
 Qy 357 KAPDMQGYFTYVLLFTYACIQTLYLHOYFHIQVYSGKRIKTAIVIGAVYKALVITNSARK 416
 Db 357 KAPDMQGYFTYVLLFTYACIQTLYLHOYFHIQVYSGKRIKTAIVIGAVYKALVITNSARK 416
 Qy 417 SSTVGEIVINMSVDAORFMDLATYINMISAPLOVLTALVLMINGPSVLGAVNVLM 476
 Db 417 SSTVGEIVINMSVDAORFMDLATYINMISAPLOVLTALVLMINGPSVLGAVNVLM 476
 Qy 476 VPIVAVAMAKTKTYQVAVHMSKONRIKLMNLIINGIKVLKLYAMELAFKQVLAIRQBEL 536
 Db 476 VPIVAVAMAKTKTYQVAVHMSKONRIKLMNLIINGIKVLKLYAMELAFKQVLAIRQBEL 536
 Qy 537 KVLKKSAYLSAVGFTVCTPPLVALCTPAVYVTTIDENNITLDAQAFVSLAFNLRFPL 596
 Db 537 KVLKKSAYLSAVGFTVCTPPLVALCTPAVYVTTIDENNITLDAQAFVSLAFNLRFPL 596
 Qy 597 NILPMYSSIVQASVSLKRLRFLSHHELEPDSIERPVVDGGGTNSITVPMATFTARS 656
 Db 597 NILPMYSSIVQASVSLKRLRFLSHHELEPDSIERPVVDGGGTNSITVPMATFTARS 656
 Qy 657 DPTLNGITFISIEGALVAVVGVGCGKSSLLSALLAEMDKVGHVAKISVAVVPOQAM 716
 Db 657 DPTLNGITFISIEGALVAVVGVGCGKSSLLSALLAEMDKVGHVAKISVAVVPOQAM 716
 Qy 716 IONSLSRENILFSCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEGVNLGGQKORV 776
 Db 716 IONSLSRENILFSCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEGVNLGGQKORV 776
 Qy 777 SLARAYSNADITLFPDPLSAVNAHVGKIHFEVIVGPKMLKXKTRILVTHSMYSYLPQVD 836
 Db 777 SLARAYSNADITLFPDPLSAVNAHVGKIHFEVIVGPKMLKXKTRILVTHSMYSYLPQVD 836
 Qy 837 VITVMSGKISIEGVSQELLARDAFAEFLRTASTAIEQODARENGTVGSGP-GKEAKQ 895
 Db 837 VITVMSGKISIEGVSQELLARDAFAEFLRTASTAIEQODARENGTVGSGP-GKEAKQ 895
 Qy 896 MEGMLVTSAGAQLOLROUSSSSYSYGD--SHHNSVTAELQAEAKGETTVMLEADKA 953
 Db 896 MEGMLVTSAGAQLOLROUSSSSYSYGD--SHHNSVTAELQAEAKGETTVMLEADKA 953
 Qy 954 QTQGVVLVYVWDYMKAIIGLFIISFLIFLFWCNHVSALASVYMLSLWTDPIVNGTQEHK 1013
 Db 954 QTQGVVLVYVWDYMKAIIGLFIISFLIFLFWCNHVSALASVYMLSLWTDPIVNGTQEHK 1013

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:25 ; Search time 39.4867 Seconds
(without alignments)
4076.547 Million cell updates/sec

Title: US-10-665-283-8

Perfect score: 10016
Sequence: 1 MALRGCASAGSDPLMDMNV.....RSVAVAKKPKFSIPDSLS 1947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RB_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	78.5	1531	1 US-08-463-092B-4	Sequence 4, Appl1
2	7860	78.5	1531	1 US-08-462-109A-4	Sequence 4, Appl1
3	7860	78.5	1531	1 US-08-460-907B-4	Sequence 4, Appl1
4	7860	78.5	1531	2 US-08-463-179A-4	Sequence 4, Appl1
5	7860	78.5	1531	2 US-08-461-384B-4	Sequence 4, Appl1
6	7860	78.5	1531	2 US-09-647-140B-19	Sequence 19, Appl1
7	7849	78.4	1531	1 US-08-141-893-2	Sequence 2, Appl1
8	7849	78.4	1531	1 US-08-463-092B-2	Sequence 2, Appl1
9	7849	78.4	1531	1 US-08-462-109A-2	Sequence 2, Appl1
10	7849	78.4	1531	1 US-08-460-907B-2	Sequence 2, Appl1
11	7849	78.4	1531	2 US-08-463-179A-2	Sequence 2, Appl1
12	7849	78.4	1531	2 US-08-461-384B-2	Sequence 2, Appl1
13	7849	78.4	1531	2 US-08-407-207A-2	Sequence 2, Appl1
14	7002.5	69.9	1528	1 US-08-463-092B-6	Sequence 6, Appl1
15	7002.5	69.9	1528	1 US-08-462-109A-6	Sequence 6, Appl1
16	7002.5	69.9	1528	1 US-08-460-907B-6	Sequence 6, Appl1
17	7002.5	69.9	1528	2 US-08-463-179A-6	Sequence 6, Appl1
18	7002.5	69.9	1528	2 US-08-461-384B-6	Sequence 6, Appl1
19	4487.5	44.8	1527	2 US-09-647-140B-6	Sequence 6, Appl1
20	4487.5	44.8	1530	2 US-09-647-140B-33	Sequence 33, Appl1
21	3391.5	33.9	1503	2 US-09-647-140B-8	Sequence 8, Appl1
22	3385.5	33.8	1503	2 US-09-792-616-3	Sequence 3, Appl1
23	3218.5	32.1	1498	2 US-09-792-616-9	Sequence 9, Appl1
24	2374.5	23.7	1621	2 US-08-972-927-3	Sequence 3, Appl1
25	2349.5	23.5	1622	2 US-08-972-927-6	Sequence 6, Appl1
26	2266.5	22.6	1325	2 US-09-647-140B-2	Sequence 2, Appl1
27	2212.5	22.1	1464	2 US-10-012-896-1008	Sequence 1008, Ap

28	2195.5	21.9	1261	2 US-09-439-313-538	Sequence 538, App
29	2195.5	21.9	1261	2 US-09-636-215-538	Sequence 538, App
30	2195.5	21.9	1261	2 US-09-685-166A-538	Sequence 538, App
31	2195.5	21.9	1261	2 US-09-679-426-538	Sequence 538, App
32	2195.5	21.9	1261	2 US-09-759-143-538	Sequence 538, App
33	2195.5	21.9	1261	2 US-09-651-236-538	Sequence 538, App
34	2195.5	21.9	1261	2 US-09-657-279-538	Sequence 538, App
35	2195.5	21.9	1261	2 US-10-012-896-538	Sequence 1009, App
36	2195.5	21.9	1261	2 US-10-012-896-1009	Sequence 1009, App
37	2118	21.1	1581	2 US-08-726-320-3	Sequence 3, Appl1
38	2118	21.1	1581	2 US-09-208-716-3	Sequence 3, Appl1
39	2115	21.1	1228	2 US-09-439-313-537	Sequence 537, App
40	2115	21.1	1228	2 US-09-636-215-537	Sequence 537, App
41	2115	21.1	1228	2 US-09-685-166A-537	Sequence 537, App
42	2115	21.1	1228	2 US-09-679-426-537	Sequence 537, App
43	2115	21.1	1228	2 US-09-759-143-537	Sequence 537, App
44	2115	21.1	1228	2 US-09-651-236-537	Sequence 537, App
45	2115	21.1	1228	2 US-09-657-279-537	Sequence 537, App

ALIGNMENTS

RESULT 1
US-08-463-092B-4
Sequence 4, Application US/08463092B
Patent No. 5766880
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSSEE: PARTO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
City: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Streeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1531 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-463-0928-4

Query Match 78.5%; Score 7860; DB 1; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCFYLAACFPFFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCFYLAACFPFFLYLSRH 60
QY 61 DRGVIQMTPLNKTATLGFILMIYCMADLPYSPWERSRGIFLAVFLVSPILLGITLLA 120
DB 61 DRGVIQMTPLNKTATLGFILMIYCMADLPYSPWERSRGIFLAVFLVSPILLGITLLA 120
QY 121 TFLQLERRKGVSSGIMLTFMLVALVCALILRSKIMTALKBDAQVDLPFDITFYVYS 180
DB 121 TFLQLERRKGVSSGIMLTFMLVALVCALILRSKIMTALKBDAQVDLPFDITFYVYS 180
QY 181 LLILQLVLSGDSPLFSETIHDNPPCBSSASFLSRTITWTGLIVRGYRQPLBGS 240
DB 181 LLILQLVLSGDSPLFSETIHDNPPCBSSASFLSRTITWTGLIVRGYRQPLBGS 240
QY 241 LMSLNKEDTSQVYPVLYKMKKCAKTRKQPVVYVYSSKDPAQKSSKTDANEVEAL 300
DB 241 LMSLNKEDTSQVYPVLYKMKKCAKTRKQPVVYVYSSKDPAQKSSKTDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLKTPGYPFLMSFFPKALHDLMMFSGPOLIKLIFVNDTKAP 360
DB 301 IVKSPQKEMNPSLFKVLKTPGYPFLMSFFPKALHDLMMFSGPOLIKLIFVNDTKAP 360
QY 361 WQGFYTVLFTYKLCQLVLHQYFHICFVSGMKIKTAVIGAVYRKALVYNSARKSTV 420
DB 361 WQGFYTVLFTYKLCQLVLHQYFHICFVSGMKIKTAVIGAVYRKALVYNSARKSTV 420
QY 421 GEIYNLSVDQRFMDLATYINMIWSAPLOYILALYLMLNLGSPVLAVGAVMVLMPVN 480
DB 421 GEIYNLSVDQRFMDLATYINMIWSAPLOYILALYLMLNLGSPVLAVGAVMVLMPVN 480
QY 481 AVMAKTKTYOVAMHKSCKNRILKLNELLNGIKVLKYAMELAFKDYLAIROBELKYLK 540
DB 481 AVMAKTKTYOVAMHKSCKNRILKLNELLNGIKVLKYAMELAFKDYLAIROBELKYLK 540
QY 541 KSATLSAVGTFTWCTPPLVALCTPFAVYVITDENNIIDACTAFVSLALENLRPLNLP 600
DB 541 KSATLSAVGTFTWCTPPLVALCTPFAVYVITDENNIIDACTAFVSLALENLRPLNLP 600
QY 601 MVISSIVQASVSLRLRIFLSHEBELPDSIRRRPVKDGNGNSTVRNATFWARSDDPT 660
DB 601 MVISSIVQASVSLRLRIFLSHEBELPDSIRRRPVKDGNGNSTVRNATFWARSDDPT 660
QY 721 SLRENILFGCLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKNVNSGGQKRVSLAR 780
DB 721 SLRENILFGCLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKNVNSGGQKRVSLAR 780
QY 781 AVYNSNADITYLEDDPLSAVDAAVKGKIFENNVGPKMKLNKRIILVTHSMSTLPQVDVITV 840
DB 781 AVYNSNADITYLEDDPLSAVDAAVKGKIFENNVGPKMKLNKRIILVTHSMSTLPQVDVITV 840
QY 841 MSGGKISEMSGYOELLARDGAFAEFLRTYASTEODAEKNGVYGVSPGKEAKOMENG 900
DB 841 MSGGKISEMSGYOELLARDGAFAEFLRTYASTEODAEKNGVYGVSPGKEAKOMENG 900
QY 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADKAQTQVYL 960
DB 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADKAQTQVYL 960

```

```

DB 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADKAQTQVYL 960
QY 961 SVYWDYKAIGLFISFLIFLPMCNHVSALASNYMLSLMTDDPIVNGTOHTTKVRLSVYG 1020
DB 961 SVYWDYKAIGLFISFLIFLPMCNHVSALASNYMLSLMTDDPIVNGTOHTTKVRLSVYG 1020
QY 1021 ALGISQIAVFGYSMAVSIIGIILASRCLHVDLLHSILRSWSPFERTPSGNLVNRSKEL 1080
DB 1021 ALGISQIAVFGYSMAVSIIGIILASRCLHVDLLHSILRSWSPFERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPVITKFMGSLFENVIGACTVITLATAIIPPLGIYTFVQRFVYASSROL 1140
DB 1081 DTVDSMIPVITKFMGSLFENVIGACTVITLATAIIPPLGIYTFVQRFVYASSROL 1140
QY 1141 KRLSVRSRSPYSHFNELLGVSIVIRAFEOERFIHOSDLKVDENQAYPSIVANRMLA 1200
DB 1141 KRLSVRSRSPYSHFNELLGVSIVIRAFEOERFIHOSDLKVDENQAYPSIVANRMLA 1200
QY 1201 VRLKCVGNCIYLPALFAVISRHSLSAGLVLSVSYLQVTTYLNLVMSSEMETNIVA 1260
DB 1201 VRLKCVGNCIYLPALFAVISRHSLSAGLVLSVSYLQVTTYLNLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMCIQETAPSSWPQYGRVFRVYCLRYEDLDVLRHINVTNGG 1320
DB 1261 VERLKEYSETEKAPMCIQETAPSSWPQYGRVFRVYCLRYEDLDVLRHINVTNGG 1320
QY 1321 EKVGIQVGTAGKSSLTGLFRINESAGEBIIIDGINIAKIGHDLRFKTIIPQDPVL 1380
DB 1321 EKVGIQVGTAGKSSLTGLFRINESAGEBIIIDGINIAKIGHDLRFKTIIPQDPVL 1380
QY 1381 SGIIRMLDPPSOYSDDEWNTSLAHLKDFVSALPKLDHBCAEGGENSEVGRQLVCL 1440
DB 1381 SGIIRMLDPPSOYSDDEWNTSLAHLKDFVSALPKLDHBCAEGGENSEVGRQLVCL 1440
QY 1441 ARALLRRTKTLVDEAFAVDELDLLOSTITQPEPDCVLTAAHRLNTIMOTRYIVL 1500
DB 1441 ARALLRRTKTLVDEAFAVDELDLLOSTITQPEPDCVLTAAHRLNTIMOTRYIVL 1500
QY 1501 DKGIQIEYGAPSDLLQQRGLFYSMADKAGLV 1531
DB 1501 DKGIQIEYGAPSDLLQQRGLFYSMADKAGLV 1531

```

RESULT 2
 US-08-462-109A-4
 Sequence 4, Application US/08462109A
 Patent No. 5882875
 GENERAL INFORMATION:
 APPLICANT: Cole, Susan P.C.
 APPLICANT: Deeley, Roger G.
 TITLE OF INVENTION: METHODS FOR IDENTIFYING
 TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESS: LAHIVE & COCKFIELD
 STREET: 60 State Street, suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/462,109A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/966,923
 FILING DATE: 27-OCT-1992
 APPLICATION NUMBER: 08/029,340

FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-4

Query Match 78.5% Score 7860 DB 1 Length 1531

Best Local Similarity 100.0% Pred. No. 0 Mismatches 0 Indels 0 Gaps 0

Matches 1531 Conservative 0 Mismatches 0 Indels 0 Gaps 0

QY 1 MALRGFCSADGSDPLMDMNTMTNSNDFTCFQNTVLMVPCFYLMACPFYFLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTMTNSNDFTCFQNTVLMVPCFYLMACPFYFLYSRH 60

QY 61 DRGVIQMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLABVPLVSPTLGITLLA 120
DB 61 DRGVIQMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLABVPLVSPTLGITLLA 120

QY 121 TFLIQERRRGVSSGIMLTFMLVYCALAILRSKIMTAKEDAOVDLPDITFYVYFS 180
DB 121 TFLIQERRRGVSSGIMLTFMLVYCALAILRSKIMTAKEDAOVDLPDITFYVYFS 180

QY 121 TFLIQERRRGVSSGIMLTFMLVYCALAILRSKIMTAKEDAOVDLPDITFYVYFS 180
DB 121 TFLIQERRRGVSSGIMLTFMLVYCALAILRSKIMTAKEDAOVDLPDITFYVYFS 180

QY 181 LLLIQLVLSGFSDBSPLESETIHDNPNCPRESSASFLSRITFWITGILVGYRPLEGSD 240
DB 181 LLLIQLVLSGFSDBSPLESETIHDNPNCPRESSASFLSRITFWITGILVGYRPLEGSD 240

QY 241 LMSLNKEDTSEQVVPVLVKWKKECAKTRKQPVKVVYSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKWKKECAKTRKQPVKVVYSSKDPAPKSSKVDANEVEAL 300

QY 301 IVSGPOKEMNPSPFLKYLKTRGPFPLMSPPKAIHDLMSGPOILKLLKFNVDIYAPD 360
DB 301 IVSGPOKEMNPSPFLKYLKTRGPFPLMSPPKAIHDLMSGPOILKLLKFNVDIYAPD 360

QY 361 WOGFYTVLLEFVACIQLVLIHOYFHLCEVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420
DB 361 WOGFYTVLLEFVACIQLVLIHOYFHLCEVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420

QY 421 GEIYVNLMSVDAQRFMDLATYINMIWSAPLOVITLALYLLMLNLPSPVLAGVAVMLVAVN 480
DB 421 GEIYVNLMSVDAQRFMDLATYINMIWSAPLOVITLALYLLMLNLPSPVLAGVAVMLVAVN 480

QY 481 AYVAMAKTKTYOVAMHMSKDNRIKLMNEILINGIKYLLKLANELAFKDYLAIRQELVLK 540
DB 481 AYVAMAKTKTYOVAMHMSKDNRIKLMNEILINGIKYLLKLANELAFKDYLAIRQELVLK 540

QY 541 KSAVLSAVGTFTWCTPELVALCTFAVYVITDENNIIIDAOAFVSLFNLIRPLILP 600
DB 541 KSAVLSAVGTFTWCTPELVALCTFAVYVITDENNIIIDAOAFVSLFNLIRPLILP 600

QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVNNATFTVARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVNNATFTVARSDDPT 660

QY 661 LNTGTFISIRPGALVAVVGVGCCGSSLSLALLAEMDVBEHVAIKGSVAVVPODAMQND 720
DB 661 LNTGTFISIRPGALVAVVGVGCCGSSLSLALLAEMDVBEHVAIKGSVAVVPODAMQND 720

QY 721 SLRENILFGQLEBPYRSYVIGACALLPDLIELPSGDRTEIGEGVNLSSGQQRVSILAR 780

QY 721 SLRENILFGQLEBPYRSYVIGACALLPDLIELPSGDRTEIGEGVNLSSGQQRVSILAR 780
DB 721 SLRENILFGQLEBPYRSYVIGACALLPDLIELPSGDRTEIGEGVNLSSGQQRVSILAR 780

QY 781 AVYSNADIYFPDPLSADVDAHVGKHIFENYIGPKMKLXKTRILVYHSMYLPQVDYIY 840
DB 781 AVYSNADIYFPDPLSADVDAHVGKHIFENYIGPKMKLXKTRILVYHSMYLPQVDYIY 840

QY 841 MSGKISEMSYOEILARDGAPAEFLRTVASTQEOEAENGVGVSGPKKAKOMNGM 900
DB 841 MSGKISEMSYOEILARDGAPAEFLRTVASTQEOEAENGVGVSGPKKAKOMNGM 900

QY 901 LVYDSACKQOLOROLSSSSSYSGDISRRHNSTABLOKAEKKEETKLMLEADKQGVKL 960
DB 901 LVYDSACKQOLOROLSSSSSYSGDISRRHNSTABLOKAEKKEETKLMLEADKQGVKL 960

QY 961 SVYWDYKAIQGLFISFLSIFLFCNHNVSALASNTWLSLWTDPIVNGTQHTKRLSVYG 1020
DB 961 SVYWDYKAIQGLFISFLSIFLFCNHNVSALASNTWLSLWTDPIVNGTQHTKRLSVYG 1020

QY 1021 ALGISOGIAVGYSMANVSTGILASRCLVLDLHSLRSFMSFPERPSGNLVNREKEL 1080
DB 1021 ALGISOGIAVGYSMANVSTGILASRCLVLDLHSLRSFMSFPERPSGNLVNREKEL 1080

QY 1081 DTVDSMIPVYIKAFMGSIFNVIGACIYIILATPIAIIIPPLGIYFFVORFYVASSROL 1140
DB 1081 DTVDSMIPVYIKAFMGSIFNVIGACIYIILATPIAIIIPPLGIYFFVORFYVASSROL 1140

QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEBQBRFIHQSDLKVDENQKAYYPSIVANWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEBQBRFIHQSDLKVDENQKAYYPSIVANWLA 1200

QY 1201 VRLCVCNCIVLPAALPAVYSRSLNAGLVSVSYSLQVTTIYLMVLRMSSEMETIYVA 1260
DB 1201 VRLCVCNCIVLPAALPAVYSRSLNAGLVSVSYSLQVTTIYLMVLRMSSEMETIYVA 1260

QY 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVBFNRYCARYEDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVBFNRYCARYEDLPVLRHINVTINGG 1320

QY 1321 EKVGIWGTAGKSSLLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVGIWGTAGKSSLLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380

QY 1381 SGSLRNMLDPFSQYSDBEVWTSLELAHKDPFVSLPDKLDHECAEGEENISVQQRQVCL 1440
DB 1381 SGSLRNMLDPFSQYSDBEVWTSLELAHKDPFVSLPDKLDHECAEGEENISVQQRQVCL 1440

QY 1441 ARALLKRTKILVDEATPAVDLETDDLIOSTIRTOFEDCTVLIARHANTIMDYTRYIVL 1500
DB 1441 ARALLKRTKILVDEATPAVDLETDDLIOSTIRTOFEDCTVLIARHANTIMDYTRYIVL 1500

QY 1501 DKGEIOBYGAPSDLLQORGLFYSMAKDAGIV 1531
DB 1501 DKGEIOBYGAPSDLLQORGLFYSMAKDAGIV 1531

RESULT 3
US-08-460-907B-4
Sequence 4, Application US/08460907B
Patent No. 5891724
GENERAL INFORMATION:
APPLICANT: Dealey, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
RESISTANCE ON A CELL
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: PARTRO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-4

Query March 78.5%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDNNTWNTSNPDFTCFQNTVAVWPCEFLMACFPFFELYLSRH 60
DB 1 MALRGFCSADGSDPLMDNNTWNTSNPDFTCFQNTVAVWPCEFLMACFPFFELYLSRH 60
QY 61 DRGTYQMTPLNKTATGALLMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITLLA 120
DB 61 DRGTYQMTPLNKTATGALLMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITLLA 120
QY 121 TELLOLEBRKVOSSGIMLTFMVALVGCALALIRSKITMTALKEAQNULFPDIFPYYFS 180
DB 121 TELLOLEBRKVOSSGIMLTFMVALVGCALALIRSKITMTALKEAQNULFPDIFPYYFS 180
QY 121 TELLOLEBRKVOSSGIMLTFMVALVGCALALIRSKITMTALKEAQNULFPDIFPYYFS 180
DB 121 TELLOLEBRKVOSSGIMLTFMVALVGCALALIRSKITMTALKEAQNULFPDIFPYYFS 180
QY 181 LLLIQVLVSCFSDSPLFSETIHDNPPCESSASPLSRITFMWITGLIVRGROPLESSD 240
DB 181 LLLIQVLVSCFSDSPLFSETIHDNPPCESSASPLSRITFMWITGLIVRGROPLESSD 240
QY 241 LMSINKEDTSQVVPVAVKMKKECAKTRKQPVKVVYSSKDPAPQKSSKYDANEVEAL 300
DB 241 LMSINKEDTSQVVPVAVKMKKECAKTRKQPVKVVYSSKDPAPQKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNSLPKRVLYKTGPRYPLMSFFPKAHDLMFSGPOLIKLIRFVNDTKAPD 360
DB 301 IVKSPQKEMNSLPKRVLYKTGPRYPLMSFFPKAHDLMFSGPOLIKLIRFVNDTKAPD 360
QY 361 WQGFYTYVLVFTVACLOTLYVHOYFHIQFVSGMRKTKAVIGAVYKALVITNSARKSSTV 420
DB 361 WQGFYTYVLVFTVACLOTLYVHOYFHIQFVSGMRKTKAVIGAVYKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLMLNGPSVLAVAVMVLMEVN 480

DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLMLNGPSVLAVAVMVLMEVN 480
QY 481 AVNMAKTKTYQVAMHMSKDNRIKLMEIILNGIKVLKLYAMELAFKQVLAIRQBELVYLK 540
DB 481 AVNMAKTKTYQVAMHMSKDNRIKLMEIILNGIKVLKLYAMELAFKQVLAIRQBELVYLK 540
QY 541 KSAVLSAVGFTVWCCTPFLVALCTFAVYVTIDENNIIIDAQTAVSALFNILRPNILP 600
DB 541 KSAVLSAVGFTVWCCTPFLVALCTFAVYVTIDENNIIIDAQTAVSALFNILRPNILP 600
QY 601 MVISIVQASVSLKRLAIFLSHELEPDSIERRVKDGCGTNSITVNNATFTWASDPPT 660
DB 601 MVISIVQASVSLKRLAIFLSHELEPDSIERRVKDGCGTNSITVNNATFTWASDPPT 660
QY 661 LMGITFSIPEGALVAVVQVCGCKSSILSLALMEQVBEHVALKGSVAIVPOQAWIQND 720
DB 661 LMGITFSIPEGALVAVVQVCGCKSSILSLALMEQVBEHVALKGSVAIVPOQAWIQND 720
QY 721 SLRENILFGQLEBPYRSVYQACALLPDLEILPSGDRTEIGEXVNLSCGQKQVSLAR 780
DB 721 SLRENILFGQLEBPYRSVYQACALLPDLEILPSGDRTEIGEXVNLSCGQKQVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDAAVGHKIFENYIGPKMLKNTTRILVTHSMSYLPQVDVYIV 840
DB 781 AVYSNADIYLFDDPLSAVDAAVGHKIFENYIGPKMLKNTTRILVTHSMSYLPQVDVYIV 840
QY 841 MSGGKISEMSYQELRLARDAFAFLRTYASTEOBDAEENGVTGVSQPKAKOMENGM 900
DB 841 MSGGKISEMSYQELRLARDAFAFLRTYASTEOBDAEENGVTGVSQPKAKOMENGM 900
QY 901 LVYDSACKOLQROLSSSSSYSGDISRRHNSYABLOKAEKKEETWKLMEADKAQTGVYL 960
DB 901 LVYDSACKOLQROLSSSSSYSGDISRRHNSYABLOKAEKKEETWKLMEADKAQTGVYL 960
QY 961 SVYWDYKKAIGLPIFSLSIFLPMCNHVSALASNYWLSLWTDPLVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYKKAIGLPIFSLSIFLPMCNHVSALASNYWLSLWTDPLVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVGYSMASVIGIILASRCIAYDLHSILRSFMSPFERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVGYSMASVIGIILASRCIAYDLHSILRSFMSPFERTPSGNLVNRFSEKL 1080
QY 1081 DTVDMSIPEVIKMEGSLFNVYIGACIVILLATPIAIIIPPLGIYFPVGRFYVASSROL 1140
DB 1081 DTVDMSIPEVIKMEGSLFNVYIGACIVILLATPIAIIIPPLGIYFPVGRFYVASSROL 1140
QY 1141 KRLSVSRSPYSHFNETLIGVSYIRAFREOERTIHOSDLKVDENQKAYYPSIVANWLA 1200
DB 1141 KRLSVSRSPYSHFNETLIGVSYIRAFREOERTIHOSDLKVDENQKAYYPSIVANWLA 1200
QY 1201 VRLCEVNCIYLFALFAVISRHSLSAGLVGSYSLOVTTYINMLVRSSEKETEIVVA 1260
DB 1201 VRLCEVNCIYLFALFAVISRHSLSAGLVGSYSLOVTTYINMLVRSSEKETEIVVA 1260
QY 1261 VERLKEVSETEKEAPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHNVITNGG 1320
DB 1261 VERLKEVSETEKEAPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHNVITNGG 1320
QY 1321 EKVGIVRTGAKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGLSRMLNDPESQVSDDEWMTSLELAHLKQVSAIPKLDHECAGEGEMSVSGOROLVCL 1440
DB 1381 SGLSRMLNDPESQVSDDEWMTSLELAHLKQVSAIPKLDHECAGEGEMSVSGOROLVCL 1440
QY 1441 ABALIRKTKILVDEATAAVDLJETDDLIOSTIRFOFEDCYVLTAHRLNTIMDTTRYIVL 1500
DB 1441 ABALIRKTKILVDEATAAVDLJETDDLIOSTIRFOFEDCYVLTAHRLNTIMDTTRYIVL 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYSMAKQAGLY 1531

Db 1501 DKGEIOEYGAESDILQORGLFYMAKDAGLV 1531

RESULT 4
US-08-463-179A-4

Sequence 4, Application US/08463179A
Patent No. 6001563

GENERAL INFORMATION:

APPLICANT: Cole, Susan P. C.

APPLICANT: Deoley, Roger G.

TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street, suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,179A

FILING DATE:

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/966,923

FILING DATE: 27-OCT-1992

APPLICATION NUMBER: 08/029,340

FILING DATE: 8-MAR-1993

APPLICATION NUMBER: 08/141,893

FILING DATE: 26-OCT-1993

APPLICATION NUMBER: 08/407,207

FILING DATE: 20-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Deconti, Giulio A. Jr.

REGISTRATION NUMBER: 31,503

REFERENCE/DOCKET NUMBER: PQI-002CP8

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1531 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-463-179A-4

Query Match 78.5%; Score 7860; DB 2; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARGFSSADGSDPLMDMNTANTSNDFPKCFONTLVVPCFYLMACPFYFLYSRH 60
DB 1 MARGFSSADGSDPLMDMNTANTSNDFPKCFONTLVVPCFYLMACPFYFLYSRH 60
QY 61 DRGVIOWTPLNKTALGFLMTVCADLFYFWEBSRGFLAPVPELVSTLIGITTLA 120
DB 61 DRGVIOWTPLNKTALGFLMTVCADLFYFWEBSRGFLAPVPELVSTLIGITTLA 120
QY 121 TFLIQERRRGVSSGIMLTFWLVALVCAIILRSKIMTALKEBAQVLDPRDITFYVYS 180
DB 121 TFLIQERRRGVSSGIMLTFWLVALVCAIILRSKIMTALKEBAQVLDPRDITFYVYS 180
QY 181 LLLIQLVLSGFSRSPFLFSETHIDPNCPESSSASFLSRITFWMTGLIVNGYROPLESGD 240
DB 181 LLLIQLVLSGFSRSPFLFSETHIDPNCPESSSASFLSRITFWMTGLIVNGYROPLESGD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRQPVKVVSXKDPQPKSSKVDANEVEAL 300

Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRQPVKVVSXKDPQPKSSKVDANEVEAL 300
QY 301 IVKSPQKKNPSLFXULYKTFPGYFLMSFFPKAIHDLMSGPOILKLIKFNVDTPAD 360
Db 301 IVKSPQKKNPSLFXULYKTFPGYFLMSFFPKAIHDLMSGPOILKLIKFNVDTPAD 360
QY 361 MGGFYVVLVFWACIOTLVLHOVPHICFSGMRIRKAVI GAVYRKALVITNSARKSSTV 420
Db 361 MGGFYVVLVFWACIOTLVLHOVPHICFSGMRIRKAVI GAVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLOVILALYLMNLGVSAGVAVMLVMPVN 480
Db 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLOVILALYLMNLGVSAGVAVMLVMPVN 480
QY 481 AVNARKTKTYQVAHMSKDNRIKLMEILNGIVLKL YAWELAFKDYALIROBELKVLK 540
Db 481 AVNARKTKTYQVAHMSKDNRIKLMEILNGIVLKL YAWELAFKDYALIROBELKVLK 540
QY 541 KSAVLSAVGTFTWCPPELVALCTFPAVYVITDENNIIIDAQTAFLVSLFNLFPPLILP 600
Db 541 KSAVLSAVGTFTWCPPELVALCTFPAVYVITDENNIIIDAQTAFLVSLFNLFPPLILP 600
QY 601 MVISIYQASVSLKRLIFLSHLEBPSIERRPVKGCGGTSITVNNATFTWARSDEPT 660
Db 601 MVISIYQASVSLKRLIFLSHLEBPSIERRPVKGCGGTSITVNNATFTWARSDEPT 660
QY 661 LMGITTSIPBGALVAVVGQVCGKSSILSALLAEMDKVEGHVAKGSVAVPOAWIOND 720
Db 661 LMGITTSIPBGALVAVVGQVCGKSSILSALLAEMDKVEGHVAKGSVAVPOAWIOND 720
QY 721 SLEENILFGQLEBPYRSYIOACALLPDEIIPSGDRTIGSKVNLSCGQORVSLAR 780
Db 721 SLEENILFGQLEBPYRSYIOACALLPDEIIPSGDRTIGSKVNLSCGQORVSLAR 780
QY 781 AAVSNADIYFDDPLSADVADHVKHIFENYIGKGMKXNTRILVHMSYLVQVDYIV 840
Db 781 AAVSNADIYFDDPLSADVADHVKHIFENYIGKGMKXNTRILVHMSYLVQVDYIV 840
QY 841 MSGKISSEMSYQELLARDGAFELRTYASTEOBDAENGVGVSGPGEKAKOMENGM 900
Db 841 MSGKISSEMSYQELLARDGAFELRTYASTEOBDAENGVGVSGPGEKAKOMENGM 900
QY 901 LYTDSAGKQOLQROLSSSSSYSGDISRHNSATLQYAEAKETWKLMEADKQOTGVKL 960
Db 901 LYTDSAGKQOLQROLSSSSSYSGDISRHNSATLQYAEAKETWKLMEADKQOTGVKL 960
QY 961 SVYMDYMKALGLFISFIFLFCMNVSAALASYMLSLMTDDPIVNGTOEHTKVRLSVYG 1020
Db 961 SVYMDYMKALGLFISFIFLFCMNVSAALASYMLSLMTDDPIVNGTOEHTKVRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGILASRCLHVDLHSLIRSPMSFFERTPSGNLVNRSKEL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGILASRCLHVDLHSLIRSPMSFFERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPEYIKKFMGSLFNVIACIVILLATPIAIIIPGLIYFQRFYVASSROL 1140
Db 1081 DTVDSMIPEYIKKFMGSLFNVIACIVILLATPIAIIIPGLIYFQRFYVASSROL 1140
QY 1141 KRLSVSRSPVSHFNETLIGSVIRAFEOBQEFHOSDLKVDENKAYVPSIVANRWLA 1200
Db 1141 KRLSVSRSPVSHFNETLIGSVIRAFEOBQEFHOSDLKVDENKAYVPSIVANRWLA 1200
QY 1201 VRLCVCNCIVLPAALFAVISRHSLSAGLVGSYSLOQVTTYLVNMLVMSSEMETNIVA 1260
Db 1201 VRLCVCNCIVLPAALFAVISRHSLSAGLVGSYSLOQVTTYLVNMLVMSSEMETNIVA 1260
QY 1261 VERLKEYSSTEKAPQIOETAPSSWPQVGRFENNYCLRRREDLDFLRHINVTINGG 1320
Db 1261 VERLKEYSSTEKAPQIOETAPSSWPQVGRFENNYCLRRREDLDFLRHINVTINGG 1320
QY 1321 EKVIGVIRGAGKSSLTGLGRINESAGEIITIDGINAKIGLHDFRKITIIPPOPVLE 1380

Db 1321 EKVGIVGTGAGKSSLTIGLFRINESAGEIITIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
Qy 1381 SGLSRMLNDPFSQYSDSEEWTSLSLAIHKDFVSALPDKLDHECAEGENLSVGRQIVCL 1440
Db 1381 SGLSRMLNDPFSQYSDSEEWTSLSLAIHKDFVSALPDKLDHECAEGENLSVGRQIVCL 1440
Qy 1441 ARALLRKTIIVLDBATPAVDLETDLLIQSTIRFQEDCTVLTIAHRIINTMDYTRIVL 1500
Db 1441 ARALLRKTIIVLDBATPAVDLETDLLIQSTIRFQEDCTVLTIAHRIINTMDYTRIVL 1500
Qy 1501 DKGEIOEGAPSDLLQORGLFYMAKDGIV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYMAKDGIV 1531

RESULT 5

US-08-461-384B-4
Sequence 4, Application US/08461384B
Patent No. 6025473
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,384B
FILING DATE: 05-JUN-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01547
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-384B-4

Query Match 78.5%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALNGFCASDSDPLMDNMVNTNINPDFTKCFONTVLWVPCEFLMAQCFPEFYLSRH 60
Db 1 MALNGFCASDSDPLMDNMVNTNINPDFTKCFONTVLWVPCEFLMAQCFPEFYLSRH 60
Qy 61 DRGIQTPLNKTALGFLMLTVCWADLFYSFWEERSGIFLAPVFLVSPTLIGITTLA 120
Db 61 DRGIQTPLNKTALGFLMLTVCWADLFYSFWEERSGIFLAPVFLVSPTLIGITTLA 120

Db 61 DRGIQTPLNKTALGFLMLTVCWADLFYSFWEERSGIFLAPVFLVSPTLIGITTLA 120
Qy 121 TFLIQERRRGVSSGIMLTFMWLVALVCALAIIRSKIMTALKEBQVDFRDIIFYVYS 180
Db 121 TFLIQERRRGVSSGIMLTFMWLVALVCALAIIRSKIMTALKEBQVDFRDIIFYVYS 180
Qy 181 LLLIQVLVSCFSDRSLFSETIHDNPNCPSSSASFSLRITFMWITGLIVNGYROPLEGSD 240
Db 181 LLLIQVLVSCFSDRSLFSETIHDNPNCPSSSASFSLRITFMWITGLIVNGYROPLEGSD 240
Qy 241 LMSLNKEDTSEOVVPLVKNMKKCAKTRKOPVAVVSSDPAPKSSKVDNAEEVYAL 300
Db 241 LMSLNKEDTSEOVVPLVKNMKKCAKTRKOPVAVVSSDPAPKSSKVDNAEEVYAL 300
Qy 301 IVKSPKEMNPSEKFLVLYKTFGPYFLMSFFFKAIHDLMMFSGPOLKLLIKFVNDTRAPD 360
Db 301 IVKSPKEMNPSEKFLVLYKTFGPYFLMSFFFKAIHDLMMFSGPOLKLLIKFVNDTRAPD 360
Qy 361 WQGYFYTVLLFTVACLOTVLHOFYHICVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFYTVLLFTVACLOTVLHOFYHICVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLSVDAORPMDLATYINMTWSAPLOVITLALYLMLNGPSVLAGVAVWVLAIPVN 480
Db 421 GEIVNLSVDAORPMDLATYINMTWSAPLOVITLALYLMLNLGDSVLAGVAVWVLAIPVN 480
Qy 481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDYALIROBELKYLK 540
Db 481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDYALIROBELKYLK 540
Qy 541 KSAVLSAVGFTFWCTPPLVALCTFPANYVTIDENNIIDAOCTAPSLALFNILRPPLILP 600
Db 541 KSAVLSAVGFTFWCTPPLVALCTFPANYVTIDENNIIDAOCTAPSLALFNILRPPLILP 600
Qy 601 MVISIVQASVSLKRLIFLSHERLEPDSIERRPVKGGGNSTITVNAFTWARSDEPT 660
Db 601 MVISIVQASVSLKRLIFLSHERLEPDSIERRPVKGGGNSTITVNAFTWARSDEPT 660
Qy 661 LNCITPFIPEGALVAVVGVQVCGKSSLLSALLAEMDVEGHVAKGSVAVYPQQAIOND 720
Db 661 LNCITPFIPEGALVAVVGVQVCGKSSLLSALLAEMDVEGHVAKGSVAVYPQQAIOND 720
Qy 721 SLRENILFGCOLLEPYRYSVIOACALLPDLLEILPSGRTIEGKVNLSGGOKORVSLAR 780
Db 721 SLRENILFGCOLLEPYRYSVIOACALLPDLLEILPSGRTIEGKVNLSGGOKORVSLAR 780
Qy 781 AVYSNADIYLFDDPLSAVDAAVGHGHIENYVGPCKMKNKTRLIVTHSMSTLPQVDVYIV 840
Db 781 AVYSNADIYLFDDPLSAVDAAVGHGHIENYVGPCKMKNKTRLIVTHSMSTLPQVDVYIV 840
Qy 841 MSGKISEMGSYQELARDGAFAEFLRTYASTBOQDAENGCVTVSGPGKEAKQEMNGM 900
Db 841 MSGKISEMGSYQELARDGAFAEFLRTYASTBOQDAENGCVTVSGPGKEAKQEMNGM 900
Qy 901 LVTDASAGKOLQRLSSSSSYSGDISRHNSYTAELQKAEKKEFTWKLMEADKAOTGOVKL 960
Db 901 LVTDASAGKOLQRLSSSSSYSGDISRHNSYTAELQKAEKKEFTWKLMEADKAOTGOVKL 960
Qy 961 SVYWDYMKATGLFISFLSIFLFCMNHVSALASNYWLSLMTDDPIVNGTQETHKRLSYG 1020
Db 961 SVYWDYMKATGLFISFLSIFLFCMNHVSALASNYWLSLMTDDPIVNGTQETHKRLSYG 1020
Qy 1021 ALGISOGIAVFGYMAVSGIGIILASRCIHDVLDLHSILRSPMSFPEPTPSGNLVNRFSEL 1080
Db 1021 ALGISOGIAVFGYMAVSGIGIILASRCIHDVLDLHSILRSPMSFPEPTPSGNLVNRFSEL 1080
Qy 1081 DTVDMSIPEVIKMGMSLFENVIGACIVILATPIAIIIPPLGIYFFVQRFYVASSROL 1140
Db 1081 DTVDMSIPEVIKMGMSLFENVIGACIVILATPIAIIIPPLGIYFFVQRFYVASSROL 1140
Qy 1141 KRLESVSRSPVYSHFNETLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
Db 1141 KRLESVSRSPVYSHFNETLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260
DB 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDFVLRLHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDFVLRLHINVTINGG 1320
QY 1321 EKQIVERTGAGKSLTLGLFRINESABGEIILINGINAKIGLHDLFKITIIIPQDPVLF 1380
DB 1321 EKQIVERTGAGKSLTLGLFRINESABGEIILINGINAKIGLHDLFKITIIIPQDPVLF 1380
QY 1381 SGLSRMLDPPSOYSDEEWTSLLELAHKDFVSLPKLDHECAGGGENLSVGOROLVCL 1440
DB 1381 SGLSRMLDPPSOYSDEEWTSLLELAHKDFVSLPKLDHECAGGGENLSVGOROLVCL 1440
QY 1441 ARALLRKTILVLEDEATAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
DB 1441 ARALLRKTILVLEDEATAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
QY 1501 DKGEIOYGA PSDLLQORGLFYSAKADAGLV 1531
DB 1501 DKGEIOYGA PSDLLQORGLFYSAKADAGLV 1531

RESULT 6

US-09-647-140B-19
; Sequence 19, Application US/09647140B
; Patent No. 6803184
; GENERAL INFORMATION:
; APPLICANT: Fox Chase Cancer Center
; APPLICANT: Krush, Gary D.
; APPLICANT: Lee, Kun
; APPLICANT: Belinsky, Martin G.
; APPLICANT: Bain, Lisa J.
; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use Thereof
; FILE REFERENCE: PCCC 98-02
; CURRENT APPLICATION NUMBER: US/09/647,140B
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US99/06644
; PRIOR FILING DATE: 1999-03-26
; PRIOR FILING DATE: 1998-03-27
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079,759
; PRIOR APPLICATION NUMBER: 60/095,153
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-647-140B-19

Query Match 78.5%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTYNTSNPDFTKCFONTYLVWVPCFYLMACPFYFLYSR 60
DB 1 MALRGFCSADGSDPLMDMNTYNTSNPDFTKCFONTYLVWVPCFYLMACPFYFLYSR 60
QY 61 DRGYIQTPLNKTATAGFLMTIVCMADLFYFWERSRGIPLAPVFLVSPTLIGITTLA 120
DB 61 DRGYIQTPLNKTATAGFLMTIVCMADLFYFWERSRGIPLAPVFLVSPTLIGITTLA 120
QY 121 TPLIQERRKGVSGGIMLTFMWLVAVCALAIRSKIMTALKEDAVDLFRDITFYYS 180
DB 121 TPLIQERRKGVSGGIMLTFMWLVAVCALAIRSKIMTALKEDAVDLFRDITFYYS 180
QY 181 LLLIQLVLSGFSRSDSPFLSETIHDPNCPSSASFLSRITFMWITGLIVRGYQPLEGSD 240
DB 181 LLLIQLVLSGFSRSDSPFLSETIHDPNCPSSASFLSRITFMWITGLIVRGYQPLEGSD 240

DB 181 LLLIQLVLSGFSRSDSPFLSETIHDPNCPSSASFLSRITFMWITGLIVRGYQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPUYKRWKKECAKTRQPKRVVSSQDPAQPKSSYVDANEVEAL 300
DB 241 LMSLNKEDTSEOVVPUYKRWKKECAKTRQPKRVVSSQDPAQPKSSYVDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLKTFEGPYFLMSFFFKAIHDMFSGPOILKLIKFNVDKAPD 360
DB 301 IVKSPQKEMNPSLFKVLKTFEGPYFLMSFFFKAIHDMFSGPOILKLIKFNVDKAPD 360
QY 361 MGFYFTVLLFVTAQLOTLVHOYFHCIFVSGMRKIKAVIAGAVYRKALVITNSARKSTV 420
DB 361 MGFYFTVLLFVTAQLOTLVHOYFHCIFVSGMRKIKAVIAGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATTINNTWSPAPLOYIALYLTANTLGPSTAGAVNVLMPVN 480
DB 421 GEIVNLSVDAQRFMDLATTINNTWSPAPLOYIALYLTANTLGPSTAGAVNVLMPVN 480
QY 481 ANWAKTKTYQVAHMSKDNRIKLMEILNGIKVLKLYAMELAFKDKVLAIROBELKVLK 540
DB 481 ANWAKTKTYQVAHMSKDNRIKLMEILNGIKVLKLYAMELAFKDKVLAIROBELKVLK 540
QY 541 KSAVLSAVGTFTVCTPFLVALCTFAVYVITIDENNLIDAQTAVSIALFNILRPLNLP 600
DB 541 KSAVLSAVGTFTVCTPFLVALCTFAVYVITIDENNLIDAQTAVSIALFNILRPLNLP 600
QY 601 MVTSSIVQASVSLKRLRIFLSHELEPDSIERRPVKGCGGTSITTRNATFTARSPPT 660
DB 601 MVTSSIVQASVSLKRLRIFLSHELEPDSIERRPVKGCGGTSITTRNATFTARSPPT 660
QY 661 LNCITTSIPBGALVAVVGVGCGKSSLSLALMEQKVEGVAIKGSVAVPOQAWOND 720
DB 661 LNCITTSIPBGALVAVVGVGCGKSSLSLALMEQKVEGVAIKGSVAVPOQAWOND 720
QY 721 SLRENILFGQLEPEPYRSVIOACALLPDEILPSCGRTEIGKGVNLSCGQKORVSLAR 780
DB 721 SLRENILFGQLEPEPYRSVIOACALLPDEILPSCGRTEIGKGVNLSCGQKORVSLAR 780
QY 781 AVYSNADIIYFDDPLSAVDHVGKHIPENYIGKGMKNKTRILVTHSMGYLPQVDYIY 840
DB 781 AVYSNADIIYFDDPLSAVDHVGKHIPENYIGKGMKNKTRILVTHSMGYLPQVDYIY 840
QY 841 MSGGKISEMGSYOELLARDAFAEPLRTYASTOBDQDAENGVTGVSQPKKXKQWENGM 900
DB 841 MSGGKISEMGSYOELLARDAFAEPLRTYASTOBDQDAENGVTGVSQPKKXKQWENGM 900
QY 901 LVTDSAGKOLQROLSSSSYSQDISRHNSTALQKAKEETKLMLEADKQOTGOVKL 960
DB 901 LVTDSAGKOLQROLSSSSYSQDISRHNSTALQKAKEETKLMLEADKQOTGOVKL 960
QY 961 SVTYWDMKAIGLFISFLSIFLFMCNHNVSALASNYMLSLWTDPIVNGTOEHTYRLSVYG 1020
DB 961 SVTYWDMKAIGLFISFLSIFLFMCNHNVSALASNYMLSLWTDPIVNGTOEHTYRLSVYG 1020
QY 1021 ALGISOGIAVFGSMVSTGGLIASCRLVVDLHLSLRGMSFFERTPSGNLVNRSKEL 1080
DB 1021 ALGISOGIAVFGSMVSTGGLIASCRLVVDLHLSLRGMSFFERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPEYIKMFMSLFFNVIGACIVILLATPIAIIIPPLGLIYFFVORFYVASSROL 1140
DB 1081 DTVDSMIPEYIKMFMSLFFNVIGACIVILLATPIAIIIPPLGLIYFFVORFYVASSROL 1140
QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEOERFIHQSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEOERFIHQSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260
DB 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDFVLRLHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDFVLRLHINVTINGG 1320

QY	1322	EKVGIVERTAGKXSLTGLGRIMESAGEIIIDGINAKTIGLHDLRKTIIPQDPVLF	1380
	1321	EKVGIVERTAGKXSLTGLGRIMESAGEIIIDGINAKTIGLHDLRKTIIPQDPVLF	1380
QY	1381	SGSLRNMNLDPPRSQYSDEEVTWSTLELAHKDPVSALPDCLDHECAGEGNNLSVGORQVCL	1440
	1381	SGSLRNMNLDPPRSQYSDEEVTWSTLELAHKDPVSALPDCLDHECAGEGNNLSVGORQVCL	1440
QY	1441	ARALLRKTKIIVLDEATPAVDLETFDDLIQSTIRFQFEDCTVLTAAHRLNTIMDYTRVIVL	1500
	1441	ARALLRKTKIIVLDEATPAVDLETFDDLIQSTIRFQFEDCTVLTAAHRLNTIMDYTRVIVL	1500
QY	1501	DKGEIOERGASDILQORGLFYMAKXAGLV	1531
	1501	DKGEIOERGASDILQORGLFYMAKXAGLV	1531

```

Db      1201  VRLCEVNCNCTVLFALPAVLSRHSLSAGVLVSYSLSQVTTYNLVRMSSEMETIVA 1260
Qy      1261  VERLKESETEKEPMQIOETAPSSWPQVGRVFRNYCLARREDLDFVLRHINVTNGG 1320
Db      1261  VERLKESETEKEPMQIOETAPSSWPQVGRVFRNYCLARREDLDFVLRHINVTNGG 1320
Qy      1321  EKVGIVRTGAGKSSLTLGLFRINESAGEIIGININAKIGLHDLRFKTIIPQDVLV 1380
Db      1321  EKVGIVRTGAGKSSLTLGLFRINESAGEIIGININAKIGLHDLRFKTIIPQDVLV 1380
Qy      1381  SGIIRNMLDPFSQYSDSEVWTSLELAHKDFVSALPKLDHECAEGENLSVGQRLVCL 1440
Db      1381  SGIIRNMLDPFSQYSDSEVWTSLELAHKDFVSALPKLDHECAEGENLSVGQRLVCL 1440
Qy      1441  ARALLRRTKTLIVDEAFAVADLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRIVL 1500
Db      1441  ARALLRRTKTLIVDEAFAVADLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRIVL 1500
Qy      1501  DKGEIOEYCAPSDLLQORGLFYSWAKDAGLV 1531
Db      1501  DKGEIOEYCAPSDLLQORGLFYSWAKDAGLV 1531

```

```

RESULT 8
US-08-463-092B-2
; Sequence 2, Application US/08463092B
; Patent No. 5766880
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deesley, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,092B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-092B-2

Query Match      78.4%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  MALRGFCSADGSPPLMDMNTNTSNDPFCRONTLVVWPCFYLMACFPFYLSR 60
Db      1  MALRGFCSADGSPPLMDMNTNTSNDPFCRONTLVVWPCFYLMACFPFYLSR 60
Qy      61  DRGYIOWTPLNKTALGFLMWIVCAVDFYSFWSRGIPLAPVPLVSTLGIITLLA 120
Db      61  DRGYIOWTPLNKTALGFLMWIVCAVDFYSFWSRGIPLAPVPLVSTLGIITLLA 120
Qy      121  TFLIOLERRKGVSSGIMLTFWLVCAALAIRSKIMTALKEDAQVDLFRDITFYRFS 180
Db      121  TFLIOLERRKGVSSGIMLTFWLVCAALAIRSKIMTALKEDAQVDLFRDITFYRFS 180
Qy      181  LLLIOLVLSGFSRSPPLFSETIHDNPNCPRESSASPLRITFMWITGLIVRGYQPLEGSD 240
Db      181  LLLIOLVLSGFSRSPPLFSETIHDNPNCPRESSASPLRITFMWITGLIVRGYQPLEGSD 240
Qy      241  LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYVSSKDPQPKSSKYDANEVEAL 300
Db      241  LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYVSSKDPQPKSSKYDANEVEAL 300
Qy      301  IVKSPQKWNPSLPKVLVYKTPGYFLMSFFPKAIDHLMFSGPQIILKLIKPNVDTAPD 360
Db      301  IVKSPQKWNPSLPKVLVYKTPGYFLMSFFPKAIDHLMFSGPQIILKLIKPNVDTAPD 360
Qy      361  WQGYFTVLLFVYACQTLVHOYFHCFSGGRITAVGAYRALVITNSARSSIV 420
Db      361  WQGYFTVLLFVYACQTLVHOYFHCFSGGRITAVGAYRALVITNSARSSIV 420
Qy      421  GEIVNLSVDAQRFMDLATYINMISAPLOVILATYLMNLGPSVLAGVAVWLMPVN 480
Db      421  GEIVNLSVDAQRFMDLATYINMISAPLOVILATYLMNLGPSVLAGVAVWLMPVN 480
Qy      481  AVMAKTKTYQVAHMSKDNRIKLMEIINGIKVLKYAMELAFKQVLAIRQEBELVK 540
Db      481  AVMAKTKTYQVAHMSKDNRIKLMEIINGIKVLKYAMELAFKQVLAIRQEBELVK 540
Qy      541  KSAVLSAVGFTFWVCPPLVALCTFAVYVITDENNILDQTAVALPILRFPINTLP 600
Db      541  KSAVLSAVGFTFWVCPPLVALCTFAVYVITDENNILDQTAVALPILRFPINTLP 600
Qy      601  MVTSIVQASVSLKRLIFLSHELEPDSIERRPVQGGTNSITRNATFTARSDPT 660
Db      601  MVTSIVQASVSLKRLIFLSHELEPDSIERRPVQGGTNSITRNATFTARSDPT 660
Qy      661  LMGITFSIPGALVAVVGQVCGKSSLSALLAEMDKVEGHAIKGSVAVVPOQANIQND 720
Db      661  LMGITFSIPGALVAVVGQVCGKSSLSALLAEMDKVEGHAIKGSVAVVPOQANIQND 720
Qy      721  SURENITLFGQLEEPYRSVIOCALLPLELIPSDRPIEIGKGVNLSGGQQRVSLAR 780
Db      721  SURENITLFGQLEEPYRSVIOCALLPLELIPSDRPIEIGKGVNLSGGQQRVSLAR 780
Qy      781  AVYSNADITLFDPLDPAVDAHVKHIFENVIGPKGLKXKTRILTVHSHSYLPQVVIY 840
Db      781  AVYSNADITLFDPLDPAVDAHVKHIFENVIGPKGLKXKTRILTVHSHSYLPQVVIY 840
Qy      841  MSGKISEMGSYOELLARDGAPAEFLRTYASTEOEDAEENGVTGSGPKEAKQENMG 900
Db      841  MSGKISEMGSYOELLARDGAPAEFLRTYASTEOEDAEENGVTGSGPKEAKQENMG 900
Qy      901  LVTDASGKQLQRLSSSSVSGDISRHHNSTALQEAKEKERTWKLMAADAKQTOGYVL 960

```

```

Db      | 901 | LVTDSAKOIOROISSSSSSSGDLSRHNSFAEIOKAEKEEWMKMEADKAOIGQVCL 960
Qy      | 961 | SVYDYMKAIGLFTSFLSIFLFMCNHSALASNYWLSMTDPIYNGQETHKRLSYVG 1020
Db      | 961 | SVYDYMKAIGLFTSFLSIFLFMCNHSALASNYWLSMTDPIYNGQETHKRLSYVG 1020
Qy      | 1021 | ALGISOGIAVFGYSMAVISGGIILASRCIAVDLHLSILSPMSFFERTPSGNLVNRSKEL 1080
Db      | 1021 | ALGISOGIAVFGYSMAVISGGIILASRCIAVDLHLSILSPMSFFERTPSGNLVNRSKEL 1080
Qy      | 1081 | DTUDSMIPEVIKMGSLFNVIAGCIYVLLATPIAIIIPPLGLIYFVQRFYVASSRQL 1140
Db      | 1081 | DTUDSMIPEVIKMGSLFNVIAGCIYVLLATPIAIIIPPLGLIYFVQRFYVASSRQL 1140
Qy      | 1141 | KRLESVRSRPSYSHFNETILGVSIVIRAFEEQERFIHOSDLKYDENOKAYYSIVANRWLA 1200
Db      | 1141 | KRLESVRSRPSYSHFNETILGVSIVIRAFEEQERFIHOSDLKYDENOKAYYSIVANRWLA 1200
Qy      | 1201 | VRLCEVGNCLVLPALPAVISRHSLSAGLVGLSVSYSLOVTTYLNMLVRRMSSEMETNIVA 1260
Db      | 1201 | VRLCEVGNCLVLPALPAVISRHSLSAGLVGLSVSYSLOVTTYLNMLVRRMSSEMETNIVA 1260
Qy      | 1261 | VERLKEVSETEKAPWQIOETAPRESSWPQVGRVERFRNTCLRYREDLPVLHINVTINGG 1320
Db      | 1261 | VERLKEVSETEKAPWQIOETAPRESSWPQVGRVERFRNTCLRYREDLPVLHINVTINGG 1320
Qy      | 1321 | EKVGIVRTGAGKSLTGLFRINSEAGEIIIDGINAKI GLHDLRKITIIIPQDPLF 1380
Db      | 1321 | EKVGIVRTGAGKSLTGLFRINSEAGEIIIDGINAKI GLHDLRKITIIIPQDPLF 1380
Qy      | 1381 | SGLSRMLNDPSPQYSDDEEWTSLBLAHLKDFVSALPDKLDHECAEGENLSVGRQOLVCL 1440
Db      | 1381 | SGLSRMLNDPSPQYSDDEEWTSLBLAHLKDFVSALPDKLDHECAEGENLSVGRQOLVCL 1440
Qy      | 1441 | ARALLRKTIIVDEBATAVNDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDTRYIVL 1500
Db      | 1441 | ARALLRKTIIVDEBATAVNDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDTRYIVL 1500
Qy      | 1501 | DKGEIOEYGAPSDILQORGLFYMAKADGLV 1531
Db      | 1501 | DKGEIOEYGAPSDILQORGLFYMAKADGLV 1531

```

RESULT 9

US-08-462-109A-2
Sequence 2, Application US/08462109A
Patent No. 5862875

GENERAL INFORMATION:

APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,109A
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992

```

APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002CP4
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-2

```

Query Match 78.4%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy      | 1 | MALRGFSADGSDPLMDMNTWNTSNPDFTKCFONTLVLVWPCFYLMACEPFYLYSRH 60
Db      | 1 | MALRGFSADGSDPLMDMNTWNTSNPDFTKCFONTLVLVWPCFYLMACEPFYLYSRH 60
Qy      | 61 | DRGYIQTPLNKTALGFLIMVYCADLTFYSPERERGIPLAVFLVSPILGITTLLA 120
Db      | 61 | DRGYIQTPLNKTALGFLIMVYCADLTFYSPERERGIPLAVFLVSPILGITTLLA 120
Qy      | 121 | TFLQLERRRGVSGSGLMFLWVALYCALAIRSKIMTALKEBQVDLPRDITFYVYS 180
Db      | 121 | TFLQLERRRGVSGSGLMFLWVALYCALAIRSKIMTALKEBQVDLPRDITFYVYS 180
Qy      | 181 | LLLIQVLVSGFSDRSPLEFSETHDNPCEPSSASFLSRIIFWMITGLIVGRQPLEGSD 240
Db      | 181 | LLLIQVLVSGFSDRSPLEFSETHDNPCEPSSASFLSRIIFWMITGLIVGRQPLEGSD 240
Qy      | 241 | LMSLNKEDTSEQVVPVLYVKMKKECACTRKQPVYVYSSKDPAPKSSKYDANEVEAL 300
Db      | 241 | LMSLNKEDTSEQVVPVLYVKMKKECACTRKQPVYVYSSKDPAPKSSKYDANEVEAL 300
Qy      | 301 | IVKSPKEMNPSPLEKVLTKTFGPFLMSFFPKAHDLMFSGPOLKLLIKFVNDTPAD 360
Db      | 301 | IVKSPKEMNPSPLEKVLTKTFGPFLMSFFPKAHDLMFSGPOLKLLIKFVNDTPAD 360
Qy      | 361 | WQGYFYVLLFVTACLOTVLHQYFHCIVSGMRKIKTAVIGAVYRKALVITNSARKSTV 420
Db      | 361 | WQGYFYVLLFVTACLOTVLHQYFHCIVSGMRKIKTAVIGAVYRKALVITNSARKSTV 420
Qy      | 421 | GEIVNLSVDAQRPMIDATYINMISAPLOYITALLYLMLNLGSPVLAGAVALMYPVN 480
Db      | 421 | GEIVNLSVDAQRPMIDATYINMISAPLOYITALLYLMLNLGSPVLAGAVALMYPVN 480
Qy      | 481 | AVMAKTKTQVAMKSKDNRIKLMNEILNGIKYLKIYAMELAKDKVLAIRBELKYLK 540
Db      | 481 | AVMAKTKTQVAMKSKDNRIKLMNEILNGIKYLKIYAMELAKDKVLAIRBELKYLK 540
Qy      | 541 | KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNNIIDAOTAFVSLAFNIRPELNLIP 600
Db      | 541 | KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNNIIDAOTAFVSLAFNIRPELNLIP 600
Qy      | 601 | MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVNNATFTARSPPPT 660
Db      | 601 | MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVNNATFTARSPPPT 660
Qy      | 661 | LNGITPSIPGALVAVVGVGCGKSLLSALLAMDMVBEHVAIKGSVAVYPOQAWIOND 720
Db      | 661 | LNGITPSIPGALVAVVGVGCGKSLLSALLAMDMVBEHVAIKGSVAVYPOQAWIOND 720

```

```

Qy 721 SLRENILFGCLBEPYRSVIOACALLPDLLEILPSGRTEIGKGNLSGGQXORVELAR 780
    |||
Db 721 SLRENILFGCLBEPYRSVIOACALLPDLLEILPSGRTEIGKGNLSGGQXORVELAR 780
Qy 781 AVYSNADIYLPDDPLSAVDHVGKHIENVIIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
    |||
Db 781 AVYSNADIYLPDDPLSAVDHVGKHIENVIIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
Qy 841 MSGGKISEMSYOBLLARDGAFAPFLRTYASTBOBAEENGVCVSGPGKEAQMONGM 900
    |||
Db 841 MSGGKISEMSYOBLLARDGAFAPFLRTYASTBOBAEENGVCVSGPGKEAQMONGM 900
Qy 901 LVTDSAGKOLOROLSSSSSYSGDISRHHNSTAELOKAEKEEYMKMEADKATOGVYL 960
    |||
Db 901 LVTDSAGKOLOROLSSSSSYSGDISRHHNSTAELOKAEKEEYMKMEADKATOGVYL 960
Qy 961 SYVDYMKAIQGLFISPLSIFLMCNHVSALASNYMSLMTDDPIVNGTOETHKVRLSVYG 1020
    |||
Db 961 SYVDYMKAIQGLFISPLSIFLMCNHVSALASNYMSLMTDDPIVNGTOETHKVRLSVYG 1020
Qy 1021 ALGISOGIAVPGSMANVIGGIIASRCLHVDLHSLRSPMSFEERTPSGMLVNRFSKEL 1080
    |||
Db 1021 ALGISOGIAVPGSMANVIGGIIASRCLHVDLHSLRSPMSFEERTPSGMLVNRFSKEL 1080
Qy 1081 DTVDMSIPEVIKMGSLFNVIAGACIYLATPIAAIIPPLGLIYFFVORFYASSROL 1140
    |||
Db 1081 DTVDMSIPEVIKMGSLFNVIAGACIYLATPIAAIIPPLGLIYFFVORFYASSROL 1140
Qy 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEBQERFIHOSDLKVDENOKAYPSIVANRMLA 1200
    |||
Db 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEBQERFIHOSDLKVDENOKAYPSIVANRMLA 1200
Qy 1201 VRLBVCNACIVPALPAVVISRHSLSAGLVGLSVSYLOTTYANMLVRMSSEMETNIVA 1260
    |||
Db 1201 VRLBVCNACIVPALPAVVISRHSLSAGLVGLSVSYLOTTYANMLVRMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPMOIOETAPRPSMPOVGRVFRNYCLRYRBDLPVLRHINVTNGG 1320
    |||
Db 1261 VERLKEYSETEKEAPMOIOETAPRPSMPOVGRVFRNYCLRYRBDLPVLRHINVTNGG 1320
Qy 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIDIINIAKIGLHDLRFKTIIPQDPVL 1380
    |||
Db 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIDIINIAKIGLHDLRFKTIIPQDPVL 1380
Qy 1381 SSGIRPNMLDPFSQYSDDEEWTSLFLAKDFVSALPKLDHECAGEGENTSVGQRQLVCL 1440
    |||
Db 1381 SSGIRPNMLDPFSQYSDDEEWTSLFLAKDFVSALPKLDHECAGEGENTSVGQRQLVCL 1440
Qy 1441 ARALLRRTKILVDEAPAVDLETTDLIOSTIRPOFEDCTVLTAAHRLNTIMOTRYIVL 1500
    |||
Db 1441 ARALLRRTKILVDEAPAVDLETTDLIOSTIRPOFEDCTVLTAAHRLNTIMOTRYIVL 1500
Qy 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGLV 1531
    |||
Db 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGLV 1531

```

```

; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,907B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: 01551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-907B-2

Query Match 78.4%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDMNTNTNSNPDFTKCFONTLVVWPCEYLMACEPFYLYLSRH 60
    |||
Db 1 MALRGFCSADGSDPLMDMNTNTNSNPDFTKCFONTLVVWPCEYLMACEPFYLYLSRH 60
Qy 61 DRGIVQMTPLNKTKTALGFLIMIVCWADLFYSWERSRGIFLAPVPLVSTLLGITTLLA 120
    |||
Db 61 DRGIVQMTPLNKTKTALGFLIMIVCWADLFYSWERSRGIFLAPVPLVSTLLGITTLLA 120
Qy 121 TFLIOERRKGVSSGIMLTFMLVALVCALAIIRSKIMTALXBDADQVDLFRDITFYVFS 180
    |||
Db 121 TFLIOERRKGVSSGIMLTFMLVALVCALAIIRSKIMTALXBDADQVDLFRDITFYVFS 180
Qy 181 LLLIQLVLSGFSRSPFLFSETIHDPNPCPESSASFLSRITFMWITGLIIVGSRQPLEGSD 240
    |||
Db 181 LLLIQLVLSGFSRSPFLFSETIHDPNPCPESSASFLSRITFMWITGLIIVGSRQPLEGSD 240
Qy 241 LMSLNKEDTSEQVVPVIVKWKKECAKTRKQPVKYVYSSQDPAPQPESSRVYDANEVEAL 300
    |||
Db 241 LMSLNKEDTSEQVVPVIVKWKKECAKTRKQPVKYVYSSQDPAPQPESSRVYDANEVEAL 300
Qy 301 IVKSPQKEMNPSEFKLYTKTFGPFLMSFFPKAIHDLMPGSGQILKLKFNNDTKAPD 360
    |||
Db 301 IVKSPQKEMNPSEFKLYTKTFGPFLMSFFPKAIHDLMPGSGQILKLKFNNDTKAPD 360
Qy 361 WQGYFTVLLFVTAQIQTIVLHQYFHI CFVSGNRKITAVI GAAYRKALVITNSARKSSTV 420
    |||
Db 361 WQGYFTVLLFVTAQIQTIVLHQYFHI CFVSGNRKITAVI GAAYRKALVITNSARKSSTV 420

```

```

Qy 421 GEIVNLSVDAQRPMDLATYINMISAPLOVILALYLLMLNLSGSVLGAVAVMLWMPVN 480
Db 421 GEIVNLSVDAQRPMDLATYINMISAPLOVILALYLLMLNLSGSVLGAVAVMLWMPVN 480
Qy 481 ANWAMKTKYOVAMKSKDNRIKIMNELNGIKYLKIYAMELAKDKVLAIROBELVYLK 540
Db 481 ANWAMKTKYOVAMKSKDNRIKIMNELNGIKYLKIYAMELAKDKVLAIROBELVYLK 540
Qy 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTIDENNIDAOAFAPSLALFNILRPPLNLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTIDENNIDAOAFAPSLALFNILRPPLNLP 600
Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGJNSITVRNATFTWASDPPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGJNSITVRNATFTWASDPPT 660
Qy 661 LNGTHSTPBEALVAVVGVGCGKSLSLALLAEMDKVGHVALKGSVAVYVQOAMIQND 720
Db 661 LNGTHSTPBEALVAVVGVGCGKSLSLALLAEMDKVGHVALKGSVAVYVQOAMIQND 720
Qy 721 SLRENIIFGCOLBEPYRSVIOACALLPDLRILPSGRTIEGKGVNLSGGQXORVSLAR 780
Db 721 SLRENIIFGCOLBEPYRSVIOACALLPDLRILPSGRTIEGKGVNLSGGQXORVSLAR 780
Qy 781 AVYSNADIYLEDPLSAVDAAVHGKIFENVIQPKGMLKNKTRILVTHSMSTYLPQVDVILV 840
Db 781 AVYSNADIYLEDPLSAVDAAVHGKIFENVIQPKGMLKNKTRILVTHSMSTYLPQVDVILV 840
Qy 841 MSGGKISEMGSYOELLARDGAFAPFLRYASTEOBODABENGQYCVSGRGEAOMEMGM 900
Db 841 MSGGKISEMGSYOELLARDGAFAPFLRYASTEOBODABENGQYCVSGRGEAOMEMGM 900
Qy 901 LVTSAGKOLOROSSSSSSSGDISRHHNSTAELOKAKKEBTWKLMEADKAOQTGYKL 960
Db 901 LVTSAGKOLOROSSSSSSSGDISRHHNSTAELOKAKKEBTWKLMEADKAOQTGYKL 960
Qy 961 SVYWDYMKAIQLFISFLSIFLPMCNHVSALASNYWLSLMTDPIVNGTQHTKRLSVYG 1020
Db 961 SVYWDYMKAIQLFISFLSIFLPMCNHVSALASNYWLSLMTDPIVNGTQHTKRLSVYG 1020
Qy 1021 ALGISOGIAVRCGYMAVSIIGIILASRCHADVLHLSILRSPMSFPERTRSGULVNRFSKL 1080
Db 1021 ALGISOGIAVRCGYMAVSIIGIILASRCHADVLHLSILRSPMSFPERTRSGULVNRFSKL 1080
Qy 1081 DTVDSMIPEVIXKMGSLFENVIACIYVILLATPIAIIIPPLGIYFVQRFYVASSROL 1140
Db 1081 DTVDSMIPEVIXKMGSLFENVIACIYVILLATPIAIIIPPLGIYFVQRFYVASSROL 1140
Qy 1141 KRLBSVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLVKVDENOKAYPYSIVANRWLA 1200
Db 1141 KRLBSVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLVKVDENOKAYPYSIVANRWLA 1200
Qy 1201 VRLBVCVNCIYLPALPAFVIRSHSLAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
Db 1201 VRLBVCVNCIYLPALPAFVIRSHSLAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKEAPMOIQETRPPSPMPQVRVERFRNCLAKREBDLVLRHINVTINGG 1320
Db 1261 VERLKEVSETEKEAPMOIQETRPPSPMPQVRVERFRNCLAKREBDLVLRHINVTINGG 1320
Qy 1321 EKVGIIVGRTGAKGSLTLGLFRINSAAGEIIIDGINIAKIGLHDLRFKTIIPDPVLF 1380
Db 1321 EKVGIIVGRTGAKGSLTLGLFRINSAAGEIIIDGINIAKIGLHDLRFKTIIPDPVLF 1380
Qy 1381 SSGIRPMNDPSPQYSDEEVTLSLELAHLKDFVSALPDYLDHCBAGGNNLSVGORQVYCL 1440
Db 1381 SSGIRPMNDPSPQYSDEEVTLSLELAHLKDFVSALPDYLDHCBAGGNNLSVGORQVYCL 1440
Qy 1441 ABALIRKTKIIVLBATAVADLETTDDLQSTIRTRQFEDCTVLTIAHRLNTIMDYRIVVL 1500
Db 1441 ABALIRKTKIIVLBATAVADLETTDDLQSTIRTRQFEDCTVLTIAHRLNTIMDYRIVVL 1500
Qy 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531

```

```

Db 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531
RESULT 11
US-08-463-179A-2
; Sequence 2, Application US/08463179A
; Patent No. 6001563
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,179A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: P01-002CP8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-5941
; TELEFAX: (617) 227-7400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-463-179A-2
Query Match 78.4%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MALGFCGASGSDLLQORGLFYSMADAGLV 60
Db 1 MALGFCGASGSDLLQORGLFYSMADAGLV 60
Qy 61 DRGYIOMTPLNKTKTALGFLIMVCMADLFYSFERSRGIFLAVFLVSPFLIGITLLA 120
Db 61 DRGYIOMTPLNKTKTALGFLIMVCMADLFYSFERSRGIFLAVFLVSPFLIGITLLA 120
Qy 121 TFLIQLEBRKGVGSSGIMLTFMWLVAVCALAIRSKIMTALKEDAOVDLFFDITFYVFS 180
Db 121 TFLIQLEBRKGVGSSGIMLTFMWLVAVCALAIRSKIMTALKEDAOVDLFFDITFYVFS 180
Qy 181 LLLIQVLVSCFSDSPFLFSETIHDNPPCPRESSASFLRITTWMTTGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCFSDSPFLFSETIHDNPPCPRESSASFLRITTWMTTGLIVRGYRQPLEGSD 240
Qy 240 LLLIQVLVSCFSDSPFLFSETIHDNPPCPRESSASFLRITTWMTTGLIVRGYRQPLEGSD 240
Db 240 LLLIQVLVSCFSDSPFLFSETIHDNPPCPRESSASFLRITTWMTTGLIVRGYRQPLEGSD 240

```


QY 241 LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVKNVYSKDPQPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVKNVYSKDPQPKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPISLFKVLKTFPGPYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPISLFKVLKTFPGPYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPD 360
QY 361 MOGYFYVLLFVNACTOTVLHQYFHI CFSGMR IKTAVI GAVYRKALVITNSRKSTV 420
DB 361 MOGYFYVLLFVNACTOTVLHQYFHI CFSGMR IKTAVI GAVYRKALVITNSRKSTV 420
QY 421 GEIVNLMSVDAORFMDLATYINNMISAPLOVILALYLLMNTGSPVLAGVAVMYLWVYN 480
DB 421 GEIVNLMSVDAORFMDLATYINNMISAPLOVILALYLLMNTGSPVLAGVAVMYLWVYN 480
QY 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDVLAIRQELKVLK 540
DB 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDVLAIRQELKVLK 540
QY 541 KSAVLSAVGTFTWCTPPELVALCTPAVYVITDENIILDAQTAFLPILPPLNIP 600
DB 541 KSAVLSAVGTFTWCTPPELVALCTPAVYVITDENIILDAQTAFLPILPPLNIP 600
QY 601 MVISSIVQASVSLKRLIFLSHEBELBDSIERRPVKDGGTNSITVNAFTWARSPT 660
DB 601 MVISSIVQASVSLKRLIFLSHEBELBDSIERRPVKDGGTNSITVNAFTWARSPT 660
QY 661 LMGITFSIPBGALVAVVGVCCKSSLSALLAEMDRVGHVAIKGSVAVVPOQAWI 720
DB 661 LMGITFSIPBGALVAVVGVCCKSSLSALLAEMDRVGHVAIKGSVAVVPOQAWI 720
QY 721 SLFENILFGOLBEPYRYSVIOACALLPDLFILPSPGRTETGKGVNLSCGQORVSLAR 780
DB 721 SLFENILFGOLBEPYRYSVIOACALLPDLFILPSPGRTETGKGVNLSCGQORVSLAR 780
QY 781 AAVSNNDIYLPDPLSANDAVGHKIPENYIGPKMKJNTRILVTHSMGYLPOVDYIV 840
DB 781 AAVSNNDIYLPDPLSANDAVGHKIPENYIGPKMKJNTRILVTHSMGYLPOVDYIV 840
QY 841 MSGGKI SEMSGYOELLARDGAFELTYASTEOQDAEENGVTYVSGPKEAKOMENGM 900
DB 841 MSGGKI SEMSGYOELLARDGAFELTYASTEOQDAEENGVTYVSGPKEAKOMENGM 900
QY 901 LVTDSAGKOLQROLSSSSSYSGDISRRHNSTABLOKKAKEEFWKLMEDKATQGVKL 960
DB 901 LVTDSAGKOLQROLSSSSSYSGDISRRHNSTABLOKKAKEEFWKLMEDKATQGVKL 960
QY 961 SVYMDWYKAIGLFISFLIFLFCNHNVSALASNYWLSLMTDDPIVNGTOEHTKVRLSVG 1020
DB 961 SVYMDWYKAIGLFISFLIFLFCNHNVSALASNYWLSLMTDDPIVNGTOEHTKVRLSVG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLAHVLDLHLSIRSPMSFEERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLAHVLDLHLSIRSPMSFEERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMIPEYIKMPFMSLFNVIGACIVILATPIAIIIPPLGLIYFVORFYVASSROL 1140
DB 1081 DTVDSMIPEYIKMPFMSLFNVIGACIVILATPIAIIIPPLGLIYFVORFYVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNFTLLGVSVIRAFEOERFIHOSDLKVDENQCAVYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNFTLLGVSVIRAFEOERFIHOSDLKVDENQCAVYPSIVANRWLA 1200
QY 1201 VRLCVCNCIVLPAALFAVISRHSLSAGVLGSYSVLQVTTYLMLVMSSEMETNIVA 1260
DB 1201 VRLCVCNCIVLPAALFAVISRHSLSAGVLGSYSVLQVTTYLMLVMSSEMETNIVA 1260
QY 1261 VERLKESETSEKAPMOIOTETAPSSWPQVGRVFRFRYCLRYREDDLPVLRHINVTNGG 1320
DB 1261 VERLKESETSEKAPMOIOTETAPSSWPQVGRVFRFRYCLRYREDDLPVLRHINVTNGG 1320
QY 1321 EKVGIVGRTGAGKSLTLGLFRINESAGEIITIDGINIAKIGLHDLRFKTIIPQDPVLF 1380

DB 1321 EKVGIVGRTGAGKSLTLGLFRINESAGEIITIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SCSLRNMLDPFSQYSBEFVWTSLLEHLAKDFVSLADPKLDHECAEGGENLSVGQROLVCL 1440
DB 1381 SCSLRNMLDPFSQYSBEFVWTSLLEHLAKDFVSLADPKLDHECAEGGENLSVGQROLVCL 1440
QY 1441 ARALLKRTKILVDEBTAADVLETDLIGSTIRTOPEDDCTVLTAHRLNTIMDYTRYIVL 1500
DB 1441 ARALLKRTKILVDEBTAADVLETDLIGSTIRTOPEDDCTVLTAHRLNTIMDYTRYIVL 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYSMAKDAGIV 1531
DB 1501 DKGEIOEYGAPSDILQORGLFYSMAKDAGIV 1531

RESULT 12
US-08-461-384B-2
; Sequence 2, Application US/08461384B
; Patent No. 6025473
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTIQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,384B
; FILING DATE: 05-JUN-95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Mielnicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: 1linear
; MOLECULE TYPE: protein
; US-08-461-384B-2

Query Match 78.4%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGCSADSGSPILMDANVTWNTSNPDTTKCFQNTVLVWVPCFYIMACFPFFLYLSRH 60
DB 1 MALRGCSADSGSPILMDANVTWNTSNPDTTKCFQNTVLVWVPCFYIMACFPFFLYLSRH 60
QY 61 DRGYQMTPLNKTKALGLFWIVCWADLFYSFWERSRGIPLAPVFLVPTLLGITLLA 120

Db 61 DRGIQMTPLNKTKTALGFLMTVCMADLFYSFMEBSRGIFLAVFVLSPLLGITLLA 120
Qy 121 TFLIQLERRKGVSSGIMLFEMLVALCALAILRSKIMTALKEDAOYDLFRDITFYYS 180
Db 121 TFLIQLERRKGVSSGIMLFTWVALCALAILRSKIMTALKEDAOYDLFRDITFYYS 180
Qy 181 LLLIQLVLCSPDRSPLFSETHIDPNCPESSASFLSRITFMTITGLIVRGYRPLEGSD 240
Db 181 LLLIQLVLCSPDRSPLFSETHIDPNCPESSASFLSRITFMTITGLIVRGYRPLEGSD 240
Qy 241 LMSLNKEDTSEQVVPVLYKMKKECAKTRKOPVYVYSSKDPAPCKSSSKTDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLYKMKKECAKTRKOPVYVYSSKDPAPCKSSSKTDANEVEAL 300
Qy 301 IVSPKEMNPSPFLKYTFGEFELMSFFPKALHDLMMFSGPOLIKLLIFVNDTAPD 360
Db 301 IVSPKEMNPSPFLKYTFGEFELMSFFPKALHDLMMFSGPOLIKLLIFVNDTAPD 360
Qy 361 WQGYFTVLLFVTAQCLQTLVLAHQYFHCFSVGMRIKTAIVAGVYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTAQCLQTLVLAHQYFHCFSVGMRIKTAIVAGVYRKALVITNSARKSTV 420
Qy 421 GEIVNLSVDAQRMDLATTYINMTSAPLOYITLALYMLNLGSPVLAGVAVMLAMPVN 480
Db 421 GEIVNLSVDAQRMDLATTYINMTSAPLOYITLALYMLNLGSPVLAGVAVMLAMPVN 480
Qy 481 AVNMAKTKTYOVAMKSKDNRIKLMNEILNGIKYKLYAMELAFKDKVLAIOBELKYLK 540
Db 481 AVNMAKTKTYOVAMKSKDNRIKLMNEILNGIKYKLYAMELAFKDKVLAIOBELKYLK 540
Qy 541 KSAVLSAVGTFTWCTPPLVALCTFPAVYVITDENNIIDAQTAFLFNLREPLNLP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTFPAVYVITDENNIIDAQTAFLFNLREPLNLP 600
Qy 601 MVISSIYQASVSLKRLAIFLSHELEBPDSIRRPVKOGGNGNSTTNNATTTARSPPT 660
Db 601 MVISSIYQASVSLKRLAIFLSHELEBPDSIRRPVKOGGNGNSTTNNATTTARSPPT 660
Qy 661 LINGITFSIPGALVAVVGQVCGKSSLLSALLAEKDKEGHVAIKGSAVYVPOQAMIOND 720
Db 661 LINGITFSIPGALVAVVGQVCGKSSLLSALLAEKDKEGHVAIKGSAVYVPOQAMIOND 720
Qy 721 SLRENIIFGCOLLEPPYRSVIOACALLPDLILPSGRTIEGKGVNLSGGOKORVSLAR 780
Db 721 SLRENIIFGCOLLEPPYRSVIOACALLPDLILPSGRTIEGKGVNLSGGOKORVSLAR 780
Qy 781 AVYSNADIIYLFDDPLSAVDHVGKHIIFENVIGPKGMLKNKTRILVTHSMSTLPOVDYIV 840
Db 781 AVYSNADIIYLFDDPLSAVDHVGKHIIFENVIGPKGMLKNKTRILVTHSMSTLPOVDYIV 840
Qy 841 MSGGKISEMGSYQELRLARDAFAEFLRTYASTBOBODAEENGAVGVSPEAKOMENGM 900
Db 841 MSGGKISEMGSYQELRLARDAFAEFLRTYASTBOBODAEENGAVGVSPEAKOMENGM 900
Qy 901 LVTDASAKOLOROLSSSSSYSGDISRHNSSTAELQKAKEKEETWKLMEADKAOTGYKL 960
Db 901 LVTDASAKOLOROLSSSSSYSGDISRHNSSTAELQKAKEKEETWKLMEADKAOTGYKL 960
Qy 961 SVYWDYKAIQGLFISFLSIFLPMCNHVSALASNTWLSMTDDPIYNGOHTKRLSYYG 1020
Db 961 SVYWDYKAIQGLFISFLSIFLPMCNHVSALASNTWLSMTDDPIYNGOHTKRLSYYG 1020
Qy 1021 ALGISOGIAVGVMAVSIIGIILASRCIADVLHSIISPMSPFERTPSGULVNRFSKEL 1080
Db 1021 ALGISOGIAVGVMAVSIIGIILASRCIADVLHSIISPMSPFERTPSGULVNRFSKEL 1080
Qy 1081 DTVDMSIPEVIKMGSLFENYIGACTVILATPIAIIIPGLIYFVQRFVYASSROL 1140
Db 1081 DTVDMSIPEVIKMGSLFENYIGACTVILATPIAIIIPGLIYFVQRFVYASSROL 1140
Qy 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
Db 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANRWLA 1200

Db 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
Qy 1201 VRLCEVNCICIVLPAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLVRMSSEMETNIVA 1260
Db 1201 VRLCEVNCICIVLPAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLVRMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKAPMOIOETAPSSWPQYGRVBFRRNYCIRYEDDLFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPMOIOETAPSSWPQYGRVBFRRNYCIRYEDDLFVLRHINVTINGG 1320
Qy 1321 EKVGIVERTAGKSSLTGLFRINESAGEIITIDGINIAKIGLHDLAFKTIIPODVLF 1380
Db 1321 EKVGIVERTAGKSSLTGLFRINESAGEIITIDGINIAKIGLHDLAFKTIIPODVLF 1380
Qy 1381 SGLSRMNLDPFSQYSDDEWVTSLELAHLKDFVSLPDKLHECAGGENSEVGRQOLVCL 1440
Db 1381 SGLSRMNLDPFSQYSDDEWVTSLELAHLKDFVSLPDKLHECAGGENSEVGRQOLVCL 1440
Qy 1441 ARALLRRTKILVDEATAVADLETDDLIQSTIRTOFEDCTVLITAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRRTKILVDEATAVADLETDDLIQSTIRTOFEDCTVLITAHRLNTIMDYTRIVL 1500
Qy 1501 DKGEIOEYGA PSDLLOQRGLFYSMAKXAGLV 1531
Db 1501 DKGEIOEYGA PSDLLOQRGLFYSMAKXAGLV 1531

RESULT 13
US-08-407-207A-2
Sequence 2, Application US/08407207A
Patent No. 6063621
GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSER: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,207A
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01512
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-407-207A-2

Query Match 78.4% Score 7849; DB 2; Length 1531;
 Best Local Similarity 99.9%; Pred. No. 0; Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVMTNTSNDFKCFONTVLVWVPCFYLMACEPFYFLYSRH 60
 DB 1 MALRGFCSADGSDPLMDMNTVMTNTSNDFKCFONTVLVWVPCFYLMACEPFYFLYSRH 60

QY 61 DRG1QMTPLNKTALGFLMICYMADLFYSFMEBSRGIFLAFVFLSPFLIGITTLA 120
 DB 61 DRG1QMTPLNKTALGFLMICYMADLFYSFMEBSRGIFLAFVFLSPFLIGITTLA 120

QY 121 TPLIQERRRGVSSGIMLTFMVALYCALALIRSKIMTALKEADVDLPEDITFYVYFS 180
 DB 121 TPLIQERRRGVSSGIMLTFMVALYCALALIRSKIMTALKEADVDLPEDITFYVYFS 180

QY 181 LLLIQVLVSCFSDSPFLFSETIHDNPCPESSASFLSRTFEMITGLIVRGYRPLEGSD 240
 DB 181 LLLIQVLVSCFSDSPFLFSETIHDNPCPESSASFLSRTFEMITGLIVRGYRPLEGSD 240

QY 241 LMSLNKEDTSEQVVPVLVKWKKECAKTRKQPVKVVYSSKDPAPKESKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKWKKECAKTRKQPVKVVYSSKDPAPKESKVDANEVEAL 300

QY 301 IVKSPKEMNPSPFLKXLTGPFYELMSFFKATHDLMFSGPQILILKFNVDTRAPD 360
 DB 301 IVKSPKEMNPSPFLKXLTGPFYELMSFFKATHDLMFSGPQILILKFNVDTRAPD 360

QY 361 WOGFYVLLFVTACLOTVLHQYFHLCPVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420
 DB 361 WOGFYVLLFVTACLOTVLHQYFHLCPVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420

QY 421 GEIVNIMSVDAGRFMDLATYINMWSAPLOVITALLYLMLTGBSVLAGAVVWLVMPVN 480
 DB 421 GEIVNIMSVDAGRFMDLATYINMWSAPLOVITALLYLMLTGBSVLAGAVVWLVMPVN 480

QY 481 AVNAAKTKTYQVAAHMSKDNRIKLMELINGIKYLKIYANELAFKDYLAIRQELKVLK 540
 DB 481 AVNAAKTKTYQVAAHMSKDNRIKLMELINGIKYLKIYANELAFKDYLAIRQELKVLK 540

QY 541 KSAVLSAVGTFTWCTCFEVLALCFAYVYITDENNIIDAOAFASLALFILIRPELILP 600
 DB 541 KSAVLSAVGTFTWCTCFEVLALCFAYVYITDENNIIDAOAFASLALFILIRPELILP 600

QY 601 MVISSIVQSVSLKRLRIFLSHELEBDSIERRPVKDGGGNSITVYNAATFTWARSDDPT 660
 DB 601 MVISSIVQSVSLKRLRIFLSHELEBDSIERRPVKDGGGNSITVYNAATFTWARSDDPT 660

QY 720 661 LMGITTSIPREGALVAVVGVGCGSSLLSALLAMDMVEGVAIKGSVAVVPOQAWIQND 720
 DB 720 661 LMGITTSIPREGALVAVVGVGCGSSLLSALLAMDMVEGVAIKGSVAVVPOQAWIQND 720

QY 781 AVYSNMDIYLPDDPLSANDAVHGHIFENYIGPKMLKNKTRILVTHSMYILPOVDYIIV 840
 DB 781 AVYSNMDIYLPDDPLSANDAVHGHIFENYIGPKMLKNKTRILVTHSMYILPOVDYIIV 840

QY 841 MSGGKISEMGSYQELLARDGAFAPFLTYASTROEDDAEENGVYSGSGPEKAKOMENGM 900
 DB 841 MSGGKISEMGSYQELLARDGAFAPFLTYASTROEDDAEENGVYSGSGPEKAKOMENGM 900

QY 901 LVTDSAGKOROLSSSSSGDISRHHNSTALOKAKAKEETWKLMEADKXQGTQVYL 960
 DB 901 LVTDSAGKOROLSSSSSGDISRHHNSTALOKAKAKEETWKLMEADKXQGTQVYL 960

QY 961 SVYWDYKAIAGLPSISIFLFCNHNVSALASNWTSLMTDDPIVNGTQHTKVRLSVYG 1020
 DB 961 SVYWDYKAIAGLPSISIFLFCNHNVSALASNWTSLMTDDPIVNGTQHTKVRLSVYG 1020

QY 1021 ALGISOGIAVFGXSMVAGISGILLASRCLAHVDLHSLIRSPMSEFERTPSGNLVNRFSEKL 1080
 DB 1021 ALGISOGIAVFGXSMVAGISGILLASRCLAHVDLHSLIRSPMSEFERTPSGNLVNRFSEKL 1080

QY 1081 DTVDSMIPVYIKMFMSLFNVIGACIVILLATPAAIIPPLGLIYFVORFYVASSROL 1140
 DB 1081 DTVDSMIPVYIKMFMSLFNVIGACIVILLATPAAIIPPLGLIYFVORFYVASSROL 1140

QY 1141 KRLBSVRSRVSYPHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANMFLA 1200
 DB 1141 KRLBSVRSRVSYPHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANMFLA 1200

QY 1201 VRLBVCNCTVLPAAFPVAVSRHSLSAGVLSVSYSLQVTTVLMNLVRSSEMETNIVA 1260
 DB 1201 VRLBVCNCTVLPAAFPVAVSRHSLSAGVLSVSYSLQVTTVLMNLVRSSEMETNIVA 1260

QY 1261 VERLKESETEKEAPMOIQETAPSSWPQVGRVEFRNYCLUREYEDDLFVLRHINVTINGG 1320
 DB 1261 VERLKESETEKEAPMOIQETAPSSWPQVGRVEFRNYCLUREYEDDLFVLRHINVTINGG 1320

QY 1321 EKVGIVRTGAGSLLTGLFRINESAGEI11IDGINIAKIGLHDLRFKTI11PODVLV 1380
 DB 1321 EKVGIVRTGAGSLLTGLFRINESAGEI11IDGINIAKIGLHDLRFKTI11PODVLV 1380

QY 1381 SGSLRNMLDPFSGYSDSEVWTSLELAHDKDFVSALDPKLDHEGABGGENLSVQOROLVCL 1440
 DB 1381 SGSLRNMLDPFSGYSDSEVWTSLELAHDKDFVSALDPKLDHEGABGGENLSVQOROLVCL 1440

QY 1441 ARALLRKTILVDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDTYRIYL 1500
 DB 1441 ARALLRKTILVDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDTYRIYL 1500

QY 1501 DKGEIOEYGAUSDLLQORGLFYSMAKDAGLV 1531
 DB 1501 DKGEIOEYGAUSDLLQORGLFYSMAKDAGLV 1531

RESULT 14
 US-08-463-092B-6
 ; Sequence 6, Application US/08463092B
 ; Patent No. 5766880
 ; GENERAL INFORMATION:
 ; APPLICANT: Cole, Susan P.C.
 ; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PARTIQ RESEARCH & DEVELOPMENT INNOVATIONS
 ; STREET: Queen's University at Kingston
 ; CITY: Kingston
 ; STATE: Ontario
 ; COUNTRY: CANADA
 ; ZIP: K7L 3N6
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/463,092B
 ; FILING DATE: 05-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/966,923
 ; FILING DATE: 27-OCT-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/029,340
 ; FILING DATE: 8-MAR-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/141,893

FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Mienicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-092B-6

Query Match 69.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1345; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGPCSDGSDPLMDMNTWMTNTSNDFTKCFONTVLWVPCFYLMACPEFFFLYLSRH 60
DB 1 MALRSPCSADSDPLMDMNTWMTNTSNDFTKCFONTVLWVPCFYLMACPEFFFLYLSRH 60
QY 61 DRGIQNTPLNKTALGFLMTIVCMADLFYSFWERSGIGFLAPPLNLSPTLLGTTTLA 120
DB 61 DRGIQNTPLNKTALGFLMTIVCMADLFYSFWERSGIGFLAPPLNLSPTLLGTTTLA 120
QY 121 TFLIQLRRRKVOSSGIMLTFMLVALYCALILRSKIMTALKEBDAVDLFDITFPYVYS 180
DB 121 TFLIQLRRRKVOSSGIMLTFMLVALYCALILRSKIMTALKEBDAVDLFDITFPYVYS 180
QY 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRITPMWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRITPMWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSQVPPVLYGAKKCECAKTRKOPVAVYVS-SKDPAPCKSSKVDANEVEEA 299
DB 241 LMSLNKEDTSQVPPVLYGAKKCECAKTRKOPVAVYVS-SKDPAPCKSSKVDANEVEEA 299
QY 300 LIVSPKEMNPSPLEFKVLYKTGFPFLMSPEFKAIHDLAMFSGPOLIKLTKFVNDTKAP 359
DB 301 LIVSPKEMNPSPLEFKVLYKTGFPFLMSPEFKAIHDLAMFSGPOLIKLTKFVNDTKAP 359
QY 360 DMGQFYFTVLLFVTAQIOTLVLAHOYFHCIVSGMRKIKTAVI GAVYRKALVITNSARKSST 419
DB 361 DMGQFYFTVLLFVTAQIOTLVLAHOYFHCIVSGMRKIKTAVI GAVYRKALVITNSARKSST 419
QY 420 VGEIVNLSVDAQRFMDLATTYINMTWAPLOYIILALYILKTLNLPSTVLAQVAVNLVAPV 479
DB 421 VGEIVNLSVDAQRFMDLATTYINMTWAPLOYIILALYILKTLNLPSTVLAQVAVNLVAPV 479
QY 480 NAVVAMTKTYQVAMHMSKDKRIRKLMNEILNGIKVLYKYAMELAFKQDVLAIROBELKVL 539
DB 481 NAVVAMTKTYQVAMHMSKDKRIRKLMNEILNGIKVLYKYAMELAFKQDVLAIROBELKVL 539
QY 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNIILDAOTAFVSLAFNLIFPLNLT 599
DB 541 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNIILDAOTAFVSLAFNLIFPLNLT 599
QY 600 PMVTSSTIVQSVSLKRLRIFLSHEELEDSTERRRPVKGQGTNSTIVTNATFTWASRDP 659
DB 601 PMVTSSTIVQSVSLKRLRIFLSHEELEDSTERRRPVKGQGTNSTIVTNATFTWASRDP 659
QY 660 TLNGITTSIPRGALVAVVGVGCKSSLSALLAEMDKVEBHVAKSGVAVVPOQAWON 719
DB 660 TLNGITTSIPRGALVAVVGVGCKSSLSALLAEMDKVEBHVAKSGVAVVPOQAWON 719

QY 720 DSLRENTLFCQQLREPPYRSVIOACALLPDLLEILPSPGDRTEIGEGKYNLSGGQKORVSLA 779
DB 720 DSLRENTLFCQQLREPPYRSVIOACALLPDLLEILPSPGDRTEIGEGKYNLSGGQKORVSLA 779
QY 780 RAVYSNADITLFPDPLSAVDAHCKHIFENVIGKGMKXKTRLLYHNSYLPQNVIT 839
DB 780 RAVYSNADITLFPDPLSAVDAHCKHIFENVIGKGMKXKTRLLYHNSYLPQNVIT 839
QY 840 VMSGKISEMGSYOELLARDGAFAPLRTYASTEOBDAEBENGTVGSGPKKAKOMENG 899
DB 840 VMSGKISEMGSYOELLARDGAFAPLRTYASTEOBDAEBENGTVGSGPKKAKOMENG 899
QY 900 MLVTDASGKOLQROLSSSSVSGDISRHNSSTAELOAEAKKEFTWKLMEADKQIQGVK 959
DB 900 MLVTDASGKOLQROLSSSSVSGDISRHNSSTAELOAEAKKEFTWKLMEADKQIQGVK 959
QY 960 LSVYVDYMKALGLFISFLSTFLPMCHNVSLASVYMLSTMTD--PIYNGOEHKVLVSLV 1018
DB 960 LSVYVDYMKALGLFISFLSTFLPMCHNVSLASVYMLSTMTD--PIYNGOEHKVLVSLV 1018
QY 956 LSVYVNTWKALGLFISFLSTFLPMCHNVSLASVYMLSTMTD--PIYNGOEHKVLVSLV 1015
DB 956 LSVYVNTWKALGLFISFLSTFLPMCHNVSLASVYMLSTMTD--PIYNGOEHKVLVSLV 1015
QY 1019 YGALGISQGLAVFGYSMAVSIQGLASRCLHVDLHSLSPMSFFERTSPGKLVNRFPSK 1078
DB 1016 YGALGISQGLAVFGYSMAVSIQGLASRCLHVDLHSLSPMSFFERTSPGKLVNRFPSK 1075
QY 1079 ELDTVDSMTPEVITKMGSLFENVIGACTIVILATPIAIIIPYGLIYFFVQRFYVASSR 1138
DB 1076 ELDTVDSMTPEVITKMGSLFENVIGACTIVILATPIAIIIPYGLIYFFVQRFYVASSR 1135
QY 1139 QLRKLESVRSPPVYSHENETLLGVSVIRAFEDQRFTHQSDKVDENQKATPISIVANRW 1198
DB 1136 QLRKLESVRSPPVYSHENETLLGVSVIRAFEDQRFTHQSDKVDENQKATPISIVANRW 1195
QY 1199 LAVRLCEVGNCTIVLFAVLSRHSLSAGLVGHSVSYQVTTYLWMLVMSSEMETNI 1258
DB 1196 LAVRLCEVGNCTIVLFAVLSRHSLSAGLVGHSVSYQVTTYLWMLVMSSEMETNI 1255
QY 1259 VAVRLKEYSETEKAPWQIQETAPSSWPQGVGEFRNTCLARYEDLDVYLRHINVTIN 1318
DB 1256 VAVRLKEYSETEKAPWQIQETAPSSWPQGVGEFRNTCLARYEDLDVYLRHINVTIN 1315
QY 1319 GGEKVIGVGRTRGAKSSLTGLPRINSAGEEIIIDENVINAKGLHMLRFKITTIIIPDDPV 1378
DB 1316 GGEKVIGVGRTRGAKSSLTGLPRINSAGEEIIIDENVINAKGLHMLRFKITTIIIPDDPV 1375
QY 1379 LFGSLRNMNDPFGQYDEEVTSLFLAHLKDFFVSLPDKLDHECAGGENTLVGQROLV 1438
DB 1376 LFGSLRNMNDPFGQYDEEVTSLFLAHLKDFFVSLPDKLDHECAGGENTLVGQROLV 1435
QY 1439 CLARALLRKTILVLDATAVADLETDLIQSTIRIQFEDCTVLTIAHRLMTINDYRVI 1498
DB 1436 CLARALLRKTILVLDATAVADLETDLIQSTIRIQFEDCTVLTIAHRLMTINDYRVI 1495
QY 1499 VLDKGEIQEYGABSDLLQORGLFYSMAKDAGLV 1531
DB 1496 VLDKGEIQEYGABSDLLQORGLFYSMAKDAGLV 1528

RESULT 15
US-08-462-109A-6
Sequence 6, Application US/08462109A
GENERAL INFORMATION:
Applicant: Cole, Susan P. C.
Applicant: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
NUMBER OF SEQUENCES: 6
TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,109A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Decont, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQ1-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-6

Query Match 69.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFSADGSDPLDMNMTNTNSNDFKCFONTYLVWPCYYLWACPPFFLYSRH 60
DB 1 MAIRSFSDGSDPLDMNMTNTNSNDFKCFONTYLVWPCYYLWACPPFFLYSRH 60
QY 61 DRGVIQWTPUNKTKTALGFLMIVCMADLFYSFWEKSGIFLAPVLPVSPFLGITTLLA 120
DB 61 DRGVIQWTPUNKTKTALGFLMIVCMADLFYSFWEKSGIFLAPVLPVSPFLGITTLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVAVCALAIIKSKITMTALKEADQVDFRDTFFYYFS 180
DB 121 TFLIQLERRKGVSSGIMLTFWLVAVCALAIIKSKITMTALKEADQVDFRDTFFYYFS 180
QY 181 LLLIQVLSCFSDRSPFSETIHDNCPSSSASFSLRTTFWMTGLIVNGYROPLEGSD 240
DB 181 LLLIQVLSCFSDRSPFSETIHDNCPSSSASFSLRTTFWMTGLIVNGYROPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVVKMKKSCAKTRKQPVVYVS-SKDPAPQKSSKVNANEEVEA 299
DB 241 LMSLNKEDTSEOVVPLVVKMKKSCAKTRKQPVVYVS-SKDPAPQKSSKVNANEEVEA 299
QY 300 LIYKSPKEMNPSLFKLYKTFGPFYFMSFFKAIHDLMMFSGPQILKLIKFNVDTKAP 359
DB 301 LIYKSPKEMNPSLFKLYKTFGPFYFMSFFKAIHDLMMFSGPQILKLIKFNVDTKAP 359
QY 360 DMGCFYFTVLLFTTACIQTIVHQYFHIQVSGMRITTAIVIGAVYKRAVITNSAKKST 419
DB 361 DMGCFYFTVLLFTTACIQTIVHQYFHIQVSGMRITTAIVIGAVYKRAVITNSAKKST 419
QY 420 VGEIVNLSYDAORFMDLATYINMWSAPVOVILATLWLNIGPSVLAVAVWVWVMPV 479
DB 421 VGEIVNLSYDAORFMDLATYINMWSAPVOVILATLWLNIGPSVLAVAVWVWVMPV 479
QY 480 NAVVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLVYAMELAFKDKVLAIRQELKVL 539
DB 481 NAVVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLVYAMELAFKDKVLAIRQELKVL 539

QY 540 KKSAYLSAVGFTFWCTPFLVALCTPAVYVITDENNTIADAQFASLAFNILEPPLNLT 599
DB 541 KKSAYLSAVGFTFWCTPFLVALCTPAVYVITDENNTIADAQFASLAFNILEPPLNLT 600
QY 600 PMVSIIVQASVSLKRLRIFLSHELEPDSIERRPVXDGGTNSITVNNATFTWASDP 659
DB 601 PMVSIIVQASVSLKRLRIFLSHELEPDSIERRPVXDGGTNSITVNNATFTWASDP 659
QY 660 TLNGITISIPGALVAVVGVCGKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAMION 719
DB 660 TLNGITISIPGALVAVVGVCGKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAMION 719
QY 720 DSIARENITLPGQLEEPYRYSVIOACALLPDEILIPSGDRPTEIGKGVNLSSGQKQVSLA 779
DB 720 DSIARENITLPGQLEEPYRYSVIOACALLPDEILIPSGDRPTEIGKGVNLSSGQKQVSLA 779
QY 780 RAVYSNADITLFDPLSAVDAHVGKHI FENVIGPKMLKKNKTRILVTHSMNYPQVAVYI 839
DB 780 RAVYSNADITLFDPLSAVDAHVGKHI FENVIGPKMLKKNKTRILVTHSMNYPQVAVYI 839
QY 840 VMSGKISMGSTYQELLARDGAPAEFLRTYASTEQEDAEENGVTGVSGPKKAKOMENG 899
DB 840 VMSGKISMGSTYQELLARDGAPAEFLRTYASTEQEDAEENGVTGVSGPKKAKOMENG 899
QY 900 MLVTDGKQOLOROLSSSSSYSGDISRHNHSTAELOAKAKETWKLMEADKQOTGVK 959
DB 900 MLVTDGKQOLOROLSSSSSYSGDISRHNHSTAELOAKAKETWKLMEADKQOTGVK 959
QY 960 LSVYMDPKAIGLFIPLSLIFLPMCNHVSALASVYMLSTWTD--PIYNGTOEHKRVLSV 1018
DB 960 LSVYMDPKAIGLFIPLSLIFLPMCNHVSALASVYMLSTWTD--PIYNGTOEHKRVLSV 1018
QY 956 LSVYMDPKAIGLFIPLSLIFLPMCNHVSALASVYMLSTWTD--PIYNGTOEHKRVLSV 1015
DB 956 LSVYMDPKAIGLFIPLSLIFLPMCNHVSALASVYMLSTWTD--PIYNGTOEHKRVLSV 1015
QY 1019 YGALGISQGIATVGYSAVSGIGILASRCHAVDLHSLILKSPMSFFERTSPGVLNRFSSK 1078
DB 1019 YGALGISQGIATVGYSAVSGIGILASRCHAVDLHSLILKSPMSFFERTSPGVLNRFSSK 1078
QY 1016 YGALGISQGIATVGYSAVSGIGILASRCHAVDLHSLILKSPMSFFERTSPGVLNRFSSK 1075
DB 1016 YGALGISQGIATVGYSAVSGIGILASRCHAVDLHSLILKSPMSFFERTSPGVLNRFSSK 1075
QY 1079 ELDTVDSMTEPVKIMFPGSLFNVIAGCTIVLLATPILAIITPILGLIFVYQRYVASSR 1138
DB 1079 ELDTVDSMTEPVKIMFPGSLFNVIAGCTIVLLATPILAIITPILGLIFVYQRYVASSR 1138
QY 1139 QLRKLSVSRSPPYSHNEFTLGSVYRAPEBERFHOSDLKVDENQKAYPSIVANRW 1198
DB 1139 QLRKLSVSRSPPYSHNEFTLGSVYRAPEBERFHOSDLKVDENQKAYPSIVANRW 1198
QY 1199 LAVRLBCVNCIYLPALPAVISRHSLSAGLVGSYSIQVTTYLNMLVMSSEMETNI 1258
DB 1199 LAVRLBCVNCIYLPALPAVISRHSLSAGLVGSYSIQVTTYLNMLVMSSEMETNI 1258
QY 1259 VAVERLKEVSETEKAPWQIOETAPPSWPOVGRVFRNRYCLARYREDLDFVLRHINVTIN 1318
DB 1259 VAVERLKEVSETEKAPWQIOETAPPSWPOVGRVFRNRYCLARYREDLDFVLRHINVTIN 1318
QY 1319 GGEKVGIVGRTAGKSLTGLFRINESAGEIITIDGINIAKIGLHDLRKITITIPDDV 1378
DB 1319 GGEKVGIVGRTAGKSLTGLFRINESAGEIITIDGINIAKIGLHDLRKITITIPDDV 1378
QY 1379 LFSGSLRMNDLPPSOYSDSEEVNTSLBLAHLKDFVSALPDLKDEHCAEGENLTVGQROLY 1438
DB 1379 LFSGSLRMNDLPPSOYSDSEEVNTSLBLAHLKDFVSALPDLKDEHCAEGENLTVGQROLY 1438
QY 1439 CLARALLRKTIIIVDEBATVAVDLETDLIQSTIRFQEDCTVLTIAHRLNTIMDYTRVI 1498
DB 1439 CLARALLRKTIIIVDEBATVAVDLETDLIQSTIRFQEDCTVLTIAHRLNTIMDYTRVI 1498
QY 1499 VLDKGEIOEYGAAPSDDLQORGLFYSAKXAGIV 1531
DB 1499 VLDKGEIOEYGAAPSDDLQORGLFYSAKXAGIV 1531

Search completed: December 15, 2005, 15:50:04
Job time : 52.4867 secs

This Page Blank (copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:00 ; Search time 161.491 Seconds
(without alignments)
5037.528 Million cell updates/sec

Title: US-10-665-283-8

Perfect score: 10016

Sequence: 1 MALRGFCSDADGSDPLMDMNV.....RSVAVAKAKPFSPDLS 1947

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10016	100.0	1947	US-10-665-283-8	Sequence 8, Appl1
2	9873	98.6	1927	US-10-665-283-1	Sequence 1, Appl1
3	9861	98.5	1927	US-10-665-283-6	Sequence 6, Appl1
4	9704	96.9	1891	US-10-665-283-4	Sequence 4, Appl1
5	7860	78.5	1531	US-09-939-853A-86	Sequence 86, Appl1
6	7860	78.5	1531	US-10-667-891-6	Sequence 6, Appl1
7	7860	78.5	1531	US-10-807-466-6	Sequence 6, Appl1
8	7860	78.5	1531	US-10-484-577-678	Sequence 678, Appl1
9	7860	78.5	1531	US-10-889-503-19	Sequence 19, Appl1
10	7860	78.5	1531	US-10-756-149-5033	Sequence 5033, Appl1
11	7769	77.6	1515	US-09-939-853A-87	Sequence 87, Appl1
12	7763	77.3	1515	US-10-618-281-42	Sequence 42, Appl1
13	7037	70.3	1388	US-10-408-765A-1718	Sequence 1718, Appl1
14	7002	69.9	1528	US-09-939-853A-88	Sequence 88, Appl1
15	6068	60.6	1303	US-10-618-281-63	Sequence 63, Appl1
16	4836	48.3	1911	US-10-665-283-2	Sequence 2, Appl1
17	4810	48.0	1895	US-10-665-283-9	Sequence 9, Appl1
18	4667	46.6	1875	US-10-665-283-10	Sequence 10, Appl1
19	4491	44.8	1527	US-09-939-853A-83	Sequence 83, Appl1
20	4491	44.8	1527	US-09-939-853A-84	Sequence 84, Appl1
21	4491	44.8	1527	US-09-939-853A-85	Sequence 85, Appl1
22	4487	44.8	1527	US-10-889-503-6	Sequence 6, Appl1
23	4487	44.8	1527	US-10-889-503-33	Sequence 33, Appl1
24	4479	44.7	1527	US-10-618-281-55	Sequence 55, Appl1
25	4479	44.7	1527	US-09-939-853A-85	Sequence 85, Appl1
26	4396	43.9	1522	US-10-665-283-7	Sequence 7, Appl1
27	4248	42.4	1961	US-10-665-283-7	Sequence 7, Appl1

28	4105	41.0	1941	5	US-10-665-283-12	Sequence 12, Appl1
29	3936	39.3	1905	5	US-10-665-283-3	Sequence 3, Appl1
30	3879	38.7	1548	4	US-10-807-466-28	Sequence 28, Appl1
31	3860	38.5	1548	4	US-10-807-466-1	Sequence 1, Appl1
32	3860	38.5	1548	5	US-10-807-466-1	Sequence 1, Appl1
33	3846	38.4	1538	5	US-10-807-466-23	Sequence 23, Appl1
34	3841	38.3	1538	5	US-10-807-466-24	Sequence 24, Appl1
35	3835	38.3	1538	5	US-10-807-466-24	Sequence 24, Appl1
36	3817	38.1	1539	5	US-10-807-466-26	Sequence 26, Appl1
37	3817	38.1	1531	6	US-11-097-143-13266	Sequence 13266, A
38	3797	37.9	1538	5	US-10-807-466-27	Sequence 27, Appl1
39	3779	37.7	1538	5	US-10-807-466-16	Sequence 16, Appl1
40	3774	37.7	1538	5	US-10-807-466-16	Sequence 16, Appl1
41	3773	37.7	1538	5	US-10-807-466-22	Sequence 22, Appl1
42	3768	37.6	1538	5	US-10-807-466-17	Sequence 17, Appl1
43	3754	37.5	1538	5	US-10-807-466-19	Sequence 19, Appl1
44	3750	37.4	1539	5	US-10-807-466-20	Sequence 20, Appl1
45	3733	37.3	1547	6	US-11-097-143-25248	Sequence 25248, A

ALIGNMENTS

RESULT 1									
US-10-665-283-8									
Sequence 8, Application US/10665283									
Publication No. US20050063989A1									
GENERAL INFORMATION:									
APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE									
APPLICANT: DERAND, Renaud									
APPLICANT: GARCIA, Elisabeth									
APPLICANT: PROST, Anne-Lise									
APPLICANT: REVILLAUD, Jean									
APPLICANT: VIVAUDOU, Michel									
TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREO									
FILE REFERENCE: F263US104									
CURRENT APPLICATION NUMBER: US/10/665,283									
CURRENT FILING DATE: 2003-09-22									
NUMBER OF SEQ ID NOS: 25									
SOFTWARE: PatentIn version 3.1									
LENGTH: 1947									
SEQ ID NO 8									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-10-665-283-8									
Query Match									
Best Local Similarity 100.0%; Score 10016; DB 5; Length 1947;									
Matches 1947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MALRGFCSDADGSDPLMDMNVNTNTPDFTKCFONTLVVWPCFYLMACPFYFLYSRH	60						
DB	1	MALRGFCSDADGSDPLMDMNVNTNTPDFTKCFONTLVVWPCFYLMACPFYFLYSRH	60						
QY	61	DRGYIOMTPLNKTKTKTALGFLMIVCWADLFYSWERSRGIFLAPVLVSLTIGITTLA	120						
DB	61	DRGYIOMTPLNKTKTKTALGFLMIVCWADLFYSWERSRGIFLAPVLVSLTIGITTLA	120						
QY	121	TFPIQERRKGVSSGIMLTFMLVALVCALAIIRSKIMTALKEADAQVDLPRDITFYVFS	180						
DB	121	TFPIQERRKGVSSGIMLTFMLVALVCALAIIRSKIMTALKEADAQVDLPRDITFYVFS	180						
QY	181	LLLIQVLVSCFSDRSPFLFSETTHDPNPCESSASFLSRITFMWITGLIVAGYRQPLEGSD	240						
DB	181	LLLIQVLVSCFSDRSPFLFSETTHDPNPCESSASFLSRITFMWITGLIVAGYRQPLEGSD	240						
QY	241	LMSLNKEDTSQVVPVLVNNWKECAKTRKOPKVVYSSKDPQOPRESSVVDANEVEAL	300						
DB	241	LMSLNKEDTSQVVPVLVNNWKECAKTRKOPKVVYSSKDPQOPRESSVVDANEVEAL	300						
QY	301	IYKSPQKNNPISLFKVLVYKTFGPFYFLMSFFFKAIHDLMMFSGPQIIKLLIKFVNDTKAPD	360						
DB	301	IYKSPQKNNPISLFKVLVYKTFGPFYFLMSFFFKAIHDLMMFSGPQIIKLLIKFVNDTKAPD	360						

```

Qy 361 WQGYFTVLLFVTAQCTLVLAHQYFHICFVSGMEIKTAIVAGVYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTAQCTLVLAHQYFHICFVSGMEIKTAIVAGVYRKALVITNSARKSTV 420
Qy 421 GEIYVNLMSVDAQRMDLATYINMTWSAPLOYITLALYMLNLGSSVLAGVAVMLAMPVN 480
Db 421 GEIYVNLMSVDAQRMDLATYINMTWSAPLOYITLALYMLNLGSSVLAGVAVMLAMPVN 480
Qy 481 AVAMAKTKTQVAMKSKDNRIKLMNEILNGIKYLUKLYAMELAPDKYLAIRQBELKYLK 540
Db 481 AVAMAKTKTQVAMKSKDNRIKLMNEILNGIKYLUKLYAMELAPDKYLAIRQBELKYLK 540
Qy 541 KSAIYASVGTFTWCTPFLVALCTPAVYVTTIDENNITLDAQTAFFSLAFNLRPEPLNLP 600
Db 541 KSAIYASVGTFTWCTPFLVALCTPAVYVTTIDENNITLDAQTAFFSLAFNLRPEPLNLP 600
Qy 601 MVTSSIVQASVSLKRLRIFLSHEBELPDSIERRPVKOGGCTNSITVRNATTTWASDPPT 660
Db 601 MVTSSIVQASVSLKRLRIFLSHEBELPDSIERRPVKOGGCTNSITVRNATTTWASDPPT 660
Qy 661 LINGITFSIPGALVAVVQVCGKSSLLSALLAEMDKYEGHVAIKGSVAVYPOQAMIOND 720
Db 661 LINGITFSIPGALVAVVQVCGKSSLLSALLAEMDKYEGHVAIKGSVAVYPOQAMIOND 720
Qy 721 SLRENIILFGCOLBEPYRSVIAQCALLPDLBILPSGRTIEIGKGVNLSGGOKORVSLAR 780
Db 721 SLRENIILFGCOLBEPYRSVIAQCALLPDLBILPSGRTIEIGKGVNLSGGOKORVSLAR 780
Qy 781 AVYENADIIYLPDDPLSAVDADVGHKIFENVIGPKMKLNKTRILVTHSMASLIPQVDVITV 840
Db 781 AVYENADIIYLPDDPLSAVDADVGHKIFENVIGPKMKLNKTRILVTHSMASLIPQVDVITV 840
Qy 841 MSGKISMGSYOELLARDGAFAPFLRTYASTBOBDAEENGVTGVSPPGKBAKOMENGM 900
Db 841 MSGKISMGSYOELLARDGAFAPFLRTYASTBOBDAEENGVTGVSPPGKBAKOMENGM 900
Qy 901 LVYDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAEKKEETWKLMEADKAQOGYKL 960
Db 901 LVYDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAEKKEETWKLMEADKAQOGYKL 960
Qy 961 SVYWDYMKALGLFISFLSIFLPMCHVSAALSNVLSLMTDDPYNGTOHTKRLSYG 1020
Db 961 SVYWDYMKALGLFISFLSIFLPMCHVSAALSNVLSLMTDDPYNGTOHTKRLSYG 1020
Qy 1021 ALGISOGIAVFGYMAVSIIGILASRCLHVDLHLSLSPSPFERTSGULVNRFSKL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGILASRCLHVDLHLSLSPSPFERTSGULVNRFSKL 1080
Qy 1081 DTVVSMIPEVIAKMFMSLFFNYIACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
Db 1081 DTVVSMIPEVIAKMFMSLFFNYIACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
Qy 1141 KRLSVSSRSVYSHNETLLGVSIVIRAFEEQERFIHQSDLKVENDOKAYPSIVANRWLA 1200
Db 1141 KRLSVSSRSVYSHNETLLGVSIVIRAFEEQERFIHQSDLKVENDOKAYPSIVANRWLA 1200
Qy 1201 VRLCVCNGCITLFAALFAVISIRHSLSAGLVGSVYSIQVTTYLNMVLRMSSEMNTIVA 1260
Db 1201 VRLCVCNGCITLFAALFAVISIRHSLSAGLVGSVYSIQVTTYLNMVLRMSSEMNTIVA 1260
Qy 1261 VERLEKEYETEKEAPWQIOETAPSPSPQVGVFEPRNTCLARYEDLDVLRHINVTINGG 1320
Db 1261 VERLEKEYETEKEAPWQIOETAPSPSPQVGVFEPRNTCLARYEDLDVLRHINVTINGG 1320
Qy 1321 EKVGIVGTGAGKSSLLGLFRINSAAGEIITIDGINIAKIGLHDLRFKRTITIPDPVLF 1380
Db 1321 EKVGIVGTGAGKSSLLGLFRINSAAGEIITIDGINIAKIGLHDLRFKRTITIPDPVLF 1380
Qy 1381 SGSLRMNLDPPSOYSDDEEVTMSLELAHLKDPVSLPDKLDHCCAGGENTLSVGQOLVCL 1440
Db 1381 SGSLRMNLDPPSOYSDDEEVTMSLELAHLKDPVSLPDKLDHCCAGGENTLSVGQOLVCL 1440

```

```

Qy 1441 ARALLRKTILVDEATAVNDLETDLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTILVDEATAVNDLETDLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRIVL 1500
Qy 1501 DKGEIOEYGA PSDLIQORGIFYSMAKADAGLVGGGGGMSLRKGIIPREYVLTIRAEPAE 1560
Db 1501 DKGEIOEYGA PSDLIQORGIFYSMAKADAGLVGGGGGMSLRKGIIPREYVLTIRAEPAE 1560
Qy 1561 PRYTRERRARRFVSKKNCNVANHKNIREQGRFLQDVFETTLVDLKPHTLLIFTMSFLCSW 1620
Db 1561 PRYTRERRARRFVSKKNCNVANHKNIREQGRFLQDVFETTLVDLKPHTLLIFTMSFLCSW 1620
Qy 1621 LLEPAMVWMLIAFAHGDLYAYMEKGITDLAPYDVPDYABGVNPCVTSIHSFSAFLF 1680
Db 1621 LLEPAMVWMLIAFAHGDLYAYMEKGITDLAPYDVPDYABGVNPCVTSIHSFSAFLF 1680
Qy 1681 SIEVQVITIGGRRVTECPALILITVQNI VGLMNAIMGCIEMKTAQHRRAETLIF 1740
Db 1681 SIEVQVITIGGRRVTECPALILITVQNI VGLMNAIMGCIEMKTAQHRRAETLIF 1740
Qy 1741 SKHAVITLRHGRLCFMLRVGDLRKSMTISATIMQVVRKTTSPGEVVPPLHQVDIPMENG 1800
Db 1741 SKHAVITLRHGRLCFMLRVGDLRKSMTISATIMQVVRKTTSPGEVVPPLHQVDIPMENG 1800
Qy 1801 VGGNGIFLVA PLIYHVIDSNSPLYDLAPSDLHHQDLBETIIVLEGVETGITTQARTS 1860
Db 1801 VGGNGIFLVA PLIYHVIDSNSPLYDLAPSDLHHQDLBETIIVLEGVETGITTQARTS 1860
Qy 1861 YLADEIIMGQRFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQLEDSDLALTLAS 1920
Db 1861 YLADEIIMGQRFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQLEDSDLALTLAS 1920
Qy 1921 SRGPLRRKRSVAVAKKPFISIPDLS 1947
Db 1921 SRGPLRRKRSVAVAKKPFISIPDLS 1947

RESULT 2
US-10-665-283-1
; Sequence 1, Application US/10665283
; Publication No. US20050063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLIOD, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: P263US104
; CURRENT APPLICATION NUMBER: US/10/665,283
; CURRENT FILING DATE: 2003-09-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1927
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-1

Query Match 98.6%; Score 9873; DB 5; Length 1927;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

```

Db 121 TFLIQLERRKGVSSGINTLFTWLVAVCALAIUKSKIMTALKEDAQVDFPDIITFYVFS 180
 Qy 181 LLLIQLVLSCSDSPFSETIHDPNCPRESSASFLSTITWMTGLIVRGYRPLEGSD 240
 Db 181 LLLIQLVLSCSDSPFSETIHDPNCPRESSASFLSTITWMTGLIVRGYRPLEGSD 240
 Qy 241 LMSLNKEDTSEQVAVYKMKKCECAKTRKQPVVYVSSKDPAPKSSKVDABEVEAL 300
 Db 241 LMSLNKEDTSEQVAVYKMKKCECAKTRKQPVVYVSSKDPAPKSSKVDABEVEAL 300
 Qy 301 IVKSPKEMNPSLFKVLKTFGPFYFLMSFPFKAIHDLMPFSGPILKLLIFVNDTYAPD 360
 Db 301 IVKSPKEMNPSLFKVLKTFGPFYFLMSFPFKAIHDLMPFSGPILKLLIFVNDTYAPD 360
 Qy 361 WQGFYFVLLFVTKLQTLVHGFHICFVSGMRKTAIVGAVTRKALVITNSARKSSTV 420
 Db 361 WQGFYFVLLFVTKLQTLVHGFHICFVSGMRKTAIVGAVTRKALVITNSARKSSTV 420
 Qy 421 GEIYNLMSVDAQRFMDLATYTNMISAPLOVYILALYILMLNGSPSLAGAVVMTLMPVN 480
 Db 421 GEIYNLMSVDAQRFMDLATYTNMISAPLOVYILALYILMLNGSPSLAGAVVMTLMPVN 480
 Qy 481 AVNANKTKTYQVAHMKSKDNRIKLANEILNGIKYIKLYAMELAFKDKVLAIROBELKVLK 540
 Db 481 AVNANKTKTYQVAHMKSKDNRIKLANEILNGIKYIKLYAMELAFKDKVLAIROBELKVLK 540
 Qy 541 KSAVLASVGTFTWCTPFLVALCTPFAVYVITDENNIIIDAQTAFYSLAFNILRPLNLP 600
 Db 541 KSAVLASVGTFTWCTPFLVALCTPFAVYVITDENNIIIDAQTAFYSLAFNILRPLNLP 600
 Qy 601 MVTSSIVQASVSLKRLAIFLSHELEPDSIERRPVQGGGNGSTTVNATPTTARSPPT 660
 Db 601 MVTSSIVQASVSLKRLAIFLSHELEPDSIERRPVQGGGNGSTTVNATPTTARSPPT 660
 Qy 661 LMGITFESIPGALVAVVGVCGCGSKSLSLALAEKMEKGVHAKGSVAVYPOQAWIOND 720
 Db 661 LMGITFESIPGALVAVVGVCGCGSKSLSLALAEKMEKGVHAKGSVAVYPOQAWIOND 720
 Qy 721 SLRENILFGCOLEBPYRSVYIACALLPDLIELPDSGRTEIGEKNVLSGGOKRVSLAR 780
 Db 721 SLRENILFGCOLEBPYRSVYIACALLPDLIELPDSGRTEIGEKNVLSGGOKRVSLAR 780
 Qy 781 AAVSNADIYLPDDELAVDAHVGHIFENYIGPKMKKNTKRIILVTHSMSTLPQOVYIV 840
 Db 781 AAVSNADIYLPDDELAVDAHVGHIFENYIGPKMKKNTKRIILVTHSMSTLPQOVYIV 840
 Qy 841 MSGGKISEMGSYOEELARDGAFAPFLRTYASTBOEODAEENGVTGVSQPKAKOMENGM 900
 Db 841 MSGGKISEMGSYOEELARDGAFAPFLRTYASTBOEODAEENGVTGVSQPKAKOMENGM 900
 Qy 901 LVTDASAGKOLOROLSSSSSYSGDISRRHNSTAELOKAEKKEETWKLMEADKAQTQVXL 960
 Db 901 LVTDASAGKOLOROLSSSSSYSGDISRRHNSTAELOKAEKKEETWKLMEADKAQTQVXL 960
 Qy 961 SYVWDVWKAIGLFTISFSLFPMCNHVSALASNTWLSLMTDDPVLNGTOEHTKRLSVYG 1020
 Db 961 SYVWDVWKAIGLFTISFSLFPMCNHVSALASNTWLSLMTDDPVLNGTOEHTKRLSVYG 1020
 Qy 1021 ALGISOGIAVFGYSMAVISGIIASRCIAVLDLHSLIRSPMSFPERIPSGNLVNRFSKEL 1080
 Db 1021 ALGISOGIAVFGYSMAVISGIIASRCIAVLDLHSLIRSPMSFPERIPSGNLVNRFSKEL 1080
 Qy 1081 DTYDSMIPBVIKPMFMSLFNVIGACIIVILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
 Db 1081 DTYDSMIPBVIKPMFMSLFNVIGACIIVILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
 Qy 1141 KRLSVSRPVYGHFNNTLGVSVIRAFEOERFIHOSDLKVDENQAYYPSVIANRWLA 1200
 Db 1141 KRLSVSRPVYGHFNNTLGVSVIRAFEOERFIHOSDLKVDENQAYYPSVIANRWLA 1200
 Qy 1201 VRLCEVNCIIVLFAALFAVISRHSLSAGVLGSYSLOQVTTYANMLVRMSSEMETNIVA 1260
 Db 1201 VRLCEVNCIIVLFAALFAVISRHSLSAGVLGSYSLOQVTTYANMLVRMSSEMETNIVA 1260

Db 1201 VRLCEVNCIIVLFAALFAVISRHSLSAGVLGSYSLOQVTTYANMLVRMSSEMETNIVA 1260
 Qy 1261 VERLKEVSETEKAPNOIOETAPSSSWPQVGRVFRNYCLARYRDDLPVLRHINVTING 1320
 Db 1261 VERLKEVSETEKAPNOIOETAPSSSWPQVGRVFRNYCLARYRDDLPVLRHINVTING 1320
 Qy 1321 EKVGIVGRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLFPKTIIPQDPLF 1380
 Db 1321 EKVGIVGRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLFPKTIIPQDPLF 1380
 Qy 1381 SGLSRNMLDPSQYSDSEVWTSLELAHKDFVSALPKDLHCEKAGEENLSVQROVLVCL 1440
 Db 1381 SGLSRNMLDPSQYSDSEVWTSLELAHKDFVSALPKDLHCEKAGEENLSVQROVLVCL 1440
 Qy 1441 ARALRRTKTLVDEATAVADLETDLIIOSTIRPOEEDCVTLTAHRLNTIMDYTRYIVL 1500
 Db 1441 ARALRRTKTLVDEATAVADLETDLIIOSTIRPOEEDCVTLTAHRLNTIMDYTRYIVL 1500
 Qy 1501 DKGEIOEYGA PSDLIQORGLFYSAKADAGVGGGGGMLSRKGIIPBEVYVTRLAEPDAE 1560
 Db 1501 DKGEIOEYGA PSDLIQORGLFYSAKADAGVGGGGGMLSRKGIIPBEVYVTRLAEPDAE 1560
 Qy 1561 PRYTRERRARFVSKKNCNVAHKNIREQGRFLQDVFTTVLDLKPHTLLIFTMSFLCSW 1620
 Db 1561 PRYTRERRARFVSKKNCNVAHKNIREQGRFLQDVFTTVLDLKPHTLLIFTMSFLCSW 1620
 Qy 1621 LLEPAMVWMLIAFAHGLIYAMEKGITDLAPYPDVDPDYAGBGTNVPCTVSIHSSSAFLF 1680
 Db 1621 LLEPAMVWMLIAFAHGLIYAMEKGITDLAPYPDVDPDYAGBGTNVPCTVSIHSSSAFLF 1680
 Qy 1681 SIEVOVITIGRGMVTECEPLATILILVQNIIVGIMINAMIGCIEMTAQAHRAEFLIF 1740
 Db 1681 SIEVOVITIGRGMVTECEPLATILILVQNIIVGIMINAMIGCIEMTAQAHRAEFLIF 1740
 Qy 1741 SKHAVITLRHGRICFMLRVDDLRSKMTISATIHQVVRKTSPEGEVAPLHOVDI PMENG 1800
 Db 1741 SKHAVITLRHGRICFMLRVDDLRSKMTISATIHQVVRKTSPEGEVAPLHOVDI PMENG 1800
 Qy 1801 VGGNGIFLVAPLIIYHVIDSNSPLYDLAPSDLHHODLEIIVILGEVETTGITTOARTS 1860
 Db 1801 VGGNGIFLVAPLIIYHVIDSNSPLYDLAPSDLHHODLEIIVILGEVETTGITTOARTS 1860
 Qy 1861 YLADELIMGORFVPIVAEEGRYSVYKSGNTIKVPTLCIARQULDEDSILDALTLAS 1920
 Db 1861 YLADELIMGORFVPIVAEEGRYSVYKSGNTIKVPTLCIARQULDEDSILDALTLAS 1920
 Qy 1921 SRGPLEKRSVAVAKAKPKFSISPDLS 1947
 Db 1901 SRGPLEKRSVAVAKAKPKFSISPDLS 1927

RESULT 3
 US-10-665-283-6
 ; Sequence 6, Application US/10665283
 ; Publication No. US20050063989A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
 ; APPLICANT: DERAND, Renaud
 ; APPLICANT: GARCIA, Bilebeth
 ; APPLICANT: PROST, Anne-lise
 ; APPLICANT: REVILLON, Jean
 ; APPLICANT: VIAUDOU, Michel
 ; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF.
 ; FILE REFERENCE: F263US104
 ; CURRENT APPLICATION NUMBER: US/10/665,283
 ; NUMBER OF SEQ. ID NOS: 25
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 1927
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-665-283-6

Query Match 98.5%; Score 9861; DB 5; Length 1927;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 20; Gaps 2;

```
Qy 1 MALRGFCSAGSDLDLMDMNTWNTSNPDFTFCQNTVLVWPCFYLMACFFPYLYLSRH 60
Db 1 MALRGFCSAGSDLDLMDMNTWNTSNPDFTFCQNTVLVWPCFYLMACFFPYLYLSRH 60
Qy 61 DRGVIQMTPLNKRTALGFLIMIVCMADLFYSFMSRSGIFLAPFLVSPILGITTLLA 120
Db 61 DRGVIQMTPLNKRTALGFLIMIVCMADLFYSFMSRSGIFLAPFLVSPILGITTLLA 120
Qy 121 TFLIQLERRKGVSSGIMLTFMWALVICALILRSKIMTALKEDAQVDLFPDITFYVYS 180
Db 121 TFLIQLERRKGVSSGIMLTFMWALVICALILRSKIMTALKEDAQVDLFPDITFYVYS 180
Qy 181 LLILQVLSCSDSPFLFSETHDPNCPRESSASFLSITFMWITGLIVRGYRQPLESD 240
Db 181 LLILQVLSCSDSPFLFSETHDPNCPRESSASFLSITFMWITGLIVRGYRQPLESD 240
Qy 241 LMSLNKEDTSEQVVPVLVKNMKCECAKTRKQPVKVYSSKDPAPKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNMKCECAKTRKQPVKVYSSKDPAPKSSKYDANEVEAL 300
Qy 301 IVKSPQKEMNPSLFKVLYKTFGPYFLMSFFPKAIDHLMFSGPOLIKLIFVNDYKAP 360
Db 301 IVKSPQKEMNPSLFKVLYKTFGPYFLMSFFPKAIDHLMFSGPOLIKLIFVNDYKAP 360
Qy 361 MQGFYVTVLFEVTCLOTLVHOYFHTCFVSGMRIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 MQGFYVTVLFEVTCLOTLVHOYFHTCFVSGMRIKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVILAYLLMLNMGPSVLGAVAVMLMPEVN 480
Db 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVILAYLLMLNMGPSVLGAVAVMLMPEVN 480
Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDYALIROBELKYLK 540
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDYALIROBELKYLK 540
Qy 541 KSAFLSVAFTFTWCTPFLVALCTPAVYVYTDENNIIIDAQTAFFSIALFNILRPNTLP 600
Db 541 KSAFLSVAFTFTWCTPFLVALCTPAVYVYTDENNIIIDAQTAFFSIALFNILRPNTLP 600
Qy 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPYKDGGSITVRNATFTMARSDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPYKDGGSITVRNATFTMARSDPT 660
Qy 661 LMGITFSIPEGALVAVVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAMIQND 720
Db 661 LMGITFSIPEGALVAVVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAMIQND 720
Qy 721 SLRENIILFCGLLEPYRSVTOACALLPDLLEILPBGRTIEGKCVNLSGGOKRVSLAR 780
Db 721 SLRENIILFCGLLEPYRSVTOACALLPDLLEILPBGRTIEGKCVNLSGGOKRVSLAR 780
Qy 781 AVYENADIYLPDDPLSAVDAAVGHKIFENYIGPKMKNKTRILVTHSMSTLPQVDVIV 840
Db 781 AVYENADIYLPDDPLSAVDAAVGHKIFENYIGPKMKNKTRILVTHSMSTLPQVDVIV 840
Qy 841 MSGGKISMGYSOELLARDGAFABFLRTYASTBOQDAEENGVTGVSFGKEAKOMENGM 900
Db 841 MSGGKISMGYSOELLARDGAFABFLRTYASTBOQDAEENGVTGVSFGKEAKOMENGM 900
Qy 901 LVTHSAGKOLOROISSSSSYSGDISRHHNSTAELOKAKKBEETKMLEADKAQTGOYKL 960
Db 901 LVTHSAGKOLOROISSSSSYSGDISRHHNSTAELOKAKKBEETKMLEADKAQTGOYKL 960
Qy 961 SVYDYMKAIGLFTSFLSIFLPMCHVASALSNVLSLMTDPTNGTQOHTKXVLSYVG 1020
Db 961 SVYDYMKAIGLFTSFLSIFLPMCHVASALSNVLSLMTDPTNGTQOHTKXVLSYVG 1020
```

```
Qy 1021 ALGISGIAVFGYSMAVSGIGIILASRCLHYDLHSILRSWSPFERPTSGULVNRFEKEL 1080
Db 1021 ALGISGIAVFGYSMAVSGIGIILASRCLHYDLHSILRSWSPFERPTSGULVNRFEKEL 1080
Qy 1081 DTVDMSIPEVIKPMGSLFNIVIGACIVILATPIAIIITPIGLIYFFVORFYAASSROL 1140
Db 1081 DTVDMSIPEVIKPMGSLFNIVIGACIVILATPIAIIITPIGLIYFFVORFYAASSROL 1140
Qy 1141 KRLESVSRSPVSHFNETLLGVSVIRAFEEOEERFIHQSDLKVDENOKAYPSIVANWMLA 1200
Db 1141 KRLESVSRSPVSHFNETLLGVSVIRAFEEOEERFIHQSDLKVDENOKAYPSIVANWMLA 1200
Qy 1201 VRLCEVNCIVLPAALFAVISRHSLSAGLVLSYSYSLOVTTYANMVRMSSEMETIVA 1260
Db 1201 VRLCEVNCIVLPAALFAVISRHSLSAGLVLSYSYSLOVTTYANMVRMSSEMETIVA 1260
Qy 1261 VERLKEYSETEKAPMGIQETAPPSSMPQVRVEFRNYCLRYRBDLPVLRIHIVTTINGG 1320
Db 1261 VERLKEYSETEKAPMGIQETAPPSSMPQVRVEFRNYCLRYRBDLPVLRIHIVTTINGG 1320
Qy 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLFKITIIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLFKITIIIPQDPVLF 1380
Qy 1381 SSGIRKMLDPSSQYSDSEEWNTSLELAHXDFVSMALPKLHCEKAGEENLSVGQROLVCL 1440
Db 1381 SSGIRKMLDPSSQYSDSEEWNTSLELAHXDFVSMALPKLHCEKAGEENLSVGQROLVCL 1440
Qy 1441 ABALIRKTKILVDEARAAVDLETTDDLIOSTIRIROPEDCTVLTAAHLANTIMDTRYIVL 1500
Db 1441 ABALIRKTKILVDEARAAVDLETTDDLIOSTIRIROPEDCTVLTAAHLANTIMDTRYIVL 1500
Qy 1501 DKGEIOEYGAPSDDLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
Db 1501 DKGEIOEYGAPSDDLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
Qy 1561 PRYTRERRARFVSKKGCNVAHKNIREQGRFLQDVFTTVLDLMPHTLLIFTMSFLCSW 1620
Db 1561 PRYTRERRARFVSKKGCNVAHKNIREQGRFLQDVFTTVLDLMPHTLLIFTMSFLCSW 1620
Qy 1621 LLFAMVMMLLAFAGDLYAMWEGITDLPAPYPVOPVYAGEGNVPCVTSIHSSSAFLF 1680
Db 1621 LLFAMVMMLLAFAGDLYAMWEGITDLPAPYPVOPVYAGEGNVPCVTSIHSSSAFLF 1680
Qy 1681 SIEVOVTIGFGRVATECEPLAILILIVONIVGIMINAMIGCIFMKTQAHRBAETLIF 1740
Db 1681 SIEVOVTIGFGRVATECEPLAILILIVONIVGIMINAMIGCIFMKTQAHRBAETLIF 1740
Qy 1741 SKPAVITLIRGRCLCFMLRVGDLKRSMTIISATIMQVVRKTTSPGEVYVPLHQVDIPMENG 1800
Db 1741 SKPAVITLIRGRCLCFMLRVGDLKRSMTIISATIMQVVRKTTSPGEVYVPLHQVDIPMENG 1800
Qy 1801 VGSNGIFLVAPLIIYHVIDSNSPLYDLAPSDLHHODEIIVILEGVEVETGITOARTS 1860
Db 1801 VGSNGIFLVAPLIIYHVIDSNSPLYDLAPSDLHHODEIIVILEGVEVETGITOARTS 1860
Qy 1861 YLADEIILMGQRFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQLDEDRSLDALTLAS 1920
Db 1861 YLADEIILMGQRFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQLDEDRSLDALTLAS 1920
Qy 1921 SRGPLKRSVAVAAKKPFSSISPDLS 1947
Db 1921 SRGPLKRSVAVAAKKPFSSISPDLS 1947
```

RESULT 4
US-10-665-283-4
; Sequence 4, Application US/10665283
; Publication No. US2005063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth

APPLICANT: PROST, Anne-Lise
APPLICANT: REVILLOU, Jean
APPLICANT: VIVAUDOU, Michel
TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
FILE REFERENCE: P26305104
CURRENT APPLICATION NUMBER: US/10/665,283
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1891
TYPE: PRT
ORGANISM: Homo sapiens
US-10-665-283-4

Query Match 96.9%; Score 9704; DB 5; Length 1891;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

QY 1 MALRGFCSADGSDPLMDMNTMTNTSNPDFTKCFONTVTVWVPCFYLMACPFYFLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTMTNTSNPDFTKCFONTVTVWVPCFYLMACPFYFLYSRH 60

QY 61 DRGYIQMTPLNKTKTALGFLIMIVCWADLFYSFWERSRGIFLAFVFLVSPILGITTLLA 120
DB 61 DRGYIQMTPLNKTKTALGFLIMIVCWADLFYSFWERSRGIFLAFVFLVSPILGITTLLA 120

QY 121 TFLIQERRKGVSSGIMLTFWLVALCALILRSKIMTALKEDAQVDFDITFYVPS 180
DB 121 TFLIQERRKGVSSGIMLTFWLVALCALILRSKIMTALKEDAQVDFDITFYVPS 180

QY 181 LLLIQVLSCFSDSPLESETHDNPCESSASFSLRTFMTGTGILVGRYRPLEGSD 240
DB 181 LLLIQVLSCFSDSPLESETHDNPCESSASFSLRTFMTGTGILVGRYRPLEGSD 240

QY 241 LMSLNKEDTSEQVPLVVKWKKECAKTRKQPVKVVYSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVPLVVKWKKECAKTRKQPVKVVYSSKDPAPKSSKVDANEVEAL 300

QY 301 IVKSPQKEMNPSLFKVLKTFGPFYFLMSFPFKALHDLMMFSGPOLIKLIFVNDTRAPD 360
DB 301 IVKSPQKEMNPSLFKVLKTFGPFYFLMSFPFKALHDLMMFSGPOLIKLIFVNDTRAPD 360

QY 361 MOGFYVVLVFTVACLOTIVLHOVPHICFVSGMRKTAIVTGAIVRKALVITNSARKSTV 420
DB 361 MOGFYVVLVFTVACLOTIVLHOVPHICFVSGMRKTAIVTGAIVRKALVITNSARKSTV 420

QY 421 GEIVNLMSVDAQREMDLATYINMIMSAPLOVITLALYLLMLNGPSVLAGVAVWVLMTPVN 480
DB 421 GEIVNLMSVDAQREMDLATYINMIMSAPLOVITLALYLLMLNGPSVLAGVAVWVLMTPVN 480

QY 481 AVNAKTKTYOVAMHMSKDNRIKLMNEILINGIKVLYAMELAFKDVLAIRBELKVLK 540
DB 481 AVNAKTKTYOVAMHMSKDNRIKLMNEILINGIKVLYAMELAFKDVLAIRBELKVLK 540

QY 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITDENNNILDAQTAFLSLFNILRPPLILP 600
DB 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITDENNNILDAQTAFLSLFNILRPPLILP 600

QY 601 MVISIVQASVSLKRLRIPLSHELEPDSTERRPVKGGGNTSTYVNAFTFARSDPT 660
DB 601 MVISIVQASVSLKRLRIPLSHELEPDSTERRPVKGGGNTSTYVNAFTFARSDPT 660

QY 661 LNGITFSIPGALVAVVGQVCGKSLLSALLAEMDVEGHVALKGSVAVYVPOQAWIOND 720
DB 661 LNGITFSIPGALVAVVGQVCGKSLLSALLAEMDVEGHVALKGSVAVYVPOQAWIOND 720

QY 721 SLRENIIFGCOLBEFFRYRYTACALPDLLETILPSGRTEIGEGVNLSSGQKORVSLAR 780
DB 721 SLRENIIFGCOLBEFFRYRYTACALPDLLETILPSGRTEIGEGVNLSSGQKORVSLAR 780

QY 781 AVYSNADIYLFDDPLASVDAVHGKIFENVIGPKMLKNKTRILVTHSMSYLPQVDVIYV 840
DB 781 AVYSNADIYLFDDPLASVDAVHGKIFENVIGPKMLKNKTRILVTHSMSYLPQVDVIYV 840

DB 781 AVYSNADIYLFDDPLASVDAVHGKIFENVIGPKMLKNKTRILVTHSMSYLPQVDVIYV 840

QY 841 MSGGKIEMSGSYQELLARDGAPAFPLRTYASTEBODAEENGVTGSGPGKEAQMENG 900
DB 841 MSGGKIEMSGSYQELLARDGAPAFPLRTYASTEBODAEENGVTGSGPGKEAQMENG 900

QY 901 LVTDSAGKOLOROLSSSSSGDLSRRHNSTAELQKAEAKKEFWKMEADKAOTGOVKL 960
DB 901 LVTDSAGKOLOROLSSSSSGDLSRRHNSTAELQKAEAKKEFWKMEADKAOTGOVKL 960

QY 961 SVYWDYMKALGLFISFLSIFLFMCNHYSAASNYMLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYMKALGLFISFLSIFLFMCNHYSAASNYMLSLMTDDPIVNGTOEHTKRLSVYG 1020

QY 1021 ALGISQIAVFGYSMAVSIQIILASRCLHYDLHSLRSRPSFEERPSGULVNRFSKEL 1080
DB 1021 ALGISQIAVFGYSMAVSIQIILASRCLHYDLHSLRSRPSFEERPSGULVNRFSKEL 1080

QY 1081 DTVDSMIPVYIKMFGSLFNVIACIVILATPIAIIIPPLGLTFEYVORFYVASSROL 1140
DB 1081 DTVDSMIPVYIKMFGSLFNVIACIVILATPIAIIIPPLGLTFEYVORFYVASSROL 1140

QY 1141 KRLESVSRSPVYSHFNETLLGVSYTRAPEEOERFIHQSDIKVDENQAYYPSIVANRWLA 1200
DB 1141 KRLESVSRSPVYSHFNETLLGVSYTRAPEEOERFIHQSDIKVDENQAYYPSIVANRWLA 1200

QY 1201 VRLCEVCNCLVPAALFAVYSRHSLSAGLVGSLVSYSLQVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVCNCLVPAALFAVYSRHSLSAGLVGSLVSYSLQVTTYINMLVRMSSEMETNIVA 1260

QY 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFRVYCYRYREDLDFVLRHNVITNGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFRVYCYRYREDLDFVLRHNVITNGG 1320

QY 1321 EKVGIVRTAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVGIVRTAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380

QY 1381 SCSLRNMLDPPQSYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGEENLSVGRQLVCL 1440
DB 1381 SCSLRNMLDPPQSYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGEENLSVGRQLVCL 1440

QY 1441 ARALLRKTKLVLDEATAVADLETDDLIOSTIRTOPEDCVLTARHLMTIMYTRIVL 1500
DB 1441 ARALLRKTKLVLDEATAVADLETDDLIOSTIRTOPEDCVLTARHLMTIMYTRIVL 1500

QY 1501 DKGEIOEYCAPSDLLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPPEYVLTRLAEDPAE 1560
DB 1501 DKGEIOEYCAPSDLLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPPEYVLTRLAEDPAE 1560

QY 1561 PRYTRERRARPVSKKNCNVAAHKNIREQRPLODVFETTLVDLKPHTLLIFWMSFLCSW 1620
DB 1561 PRYTRERRARPVSKKNCNVAAHKNIREQRPLODVFETTLVDLKPHTLLIFWMSFLCSW 1620

QY 1621 LLEFAMWMLIAFAHGDLYAEMKGIIDLAIPPYDVEDYAGEGNNVPCVTSIHSPSSAFLF 1680
DB 1621 LLEFAMWMLIAFAHGDLYAEMKGIIDLAIPPYDVEDYAGEGNNVPCVTSIHSPSSAFLF 1680

QY 1681 SIEVOVYIGFGGMYTEECPLAILIVQNIYVGINAINAMLCIFMKTAAHRAETLIF 1740
DB 1681 SIEVOVYIGFGGMYTEECPLAILIVQNIYVGINAINAMLCIFMKTAAHRAETLIF 1740

QY 1741 SKHAVITLRRHRLCFMLRVGDLKSMISATTHMOVVAKTSPGSEVPLDHOVIDPWENG 1800
DB 1741 SKHAVITLRRHRLCFMLRVGDLKSMISATTHMOVVAKTSPGSEVPLDHOVIDPWENG 1800

QY 1801 VGGNGTFIVAPLIIYVIVINSPLVDLAPSDLHHODLEIIVLLEGVETTGITTOARTS 1860
DB 1801 VGGNGTFIVAPLIIYVIVINSPLVDLAPSDLHHODLEIIVLLEGVETTGITTOARTS 1860

QY 1861 YLADEILMGRFPVPIVAEBDGRYSVDYSKRGNTIKVPTPLCTARQLEDEDS 1911
DB 1861 YLADEILMGRFPVPIVAEBDGRYSVDYSKRGNTIKVPTPLCTARQLEDEDS 1911

QY 1911 YLADEILMGRFPVPIVAEBDGRYSVDYSKRGNTIKVPTPLCTARQLEDEDS 1891
DB 1911 YLADEILMGRFPVPIVAEBDGRYSVDYSKRGNTIKVPTPLCTARQLEDEDS 1891

RESULT 5
US-09-939-853A-86
; Sequence 86, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Buxess et al.
; TITLE OF INVENTION: No. US20040039163A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-86

Query Match 78.5%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDSDSDPLMDNMTNTSNPDFTKCFONTVLVWVCEFLMACEFFYFLYSRHH 60
DB 1 MALRGFCSDSDSDPLMDNMTNTSNPDFTKCFONTVLVWVCEFLMACEFFYFLYSRHH 60
QY 61 DRGYIQTPLNKTATLALGFLIMVCMADLFYSFWRSGITLAPVFLVSPLLGTTLLA 120
DB 61 DRGYIQTPLNKTATLALGFLIMVCMADLFYSFWRSGITLAPVFLVSPLLGTTLLA 120
QY 121 TELQLERRKGVSSGIMLTFLVVALVCAALITRSKIMTALKEQOVNLFQDITFYVYFS 180
DB 121 TELQLERRKGVSSGIMLTFLVVALVCAALITRSKIMTALKEQOVNLFQDITFYVYFS 180
QY 121 TELQLERRKGVSSGIMLTFLVVALVCAALITRSKIMTALKEQOVNLFQDITFYVYFS 180
DB 121 TELQLERRKGVSSGIMLTFLVVALVCAALITRSKIMTALKEQOVNLFQDITFYVYFS 180
QY 181 LLLIQLVLSCEPDSPLFSETIHDNPPCESSASFLSRITFMWITGLIVRGYRPLESSD 240
DB 181 LLLIQLVLSCEPDSPLFSETIHDNPPCESSASFLSRITFMWITGLIVRGYRPLESSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPCKEMWNSLPFLVLYKTGRPYFLMSFFPKAIHDLMMFSGPOLIKLIFVNDTKAPD 360
DB 301 IVKSPCKEMWNSLPFLVLYKTGRPYFLMSFFPKAIHDLMMFSGPOLIKLIFVNDTKAPD 360
QY 361 WQGFYFYVLLFVTACLOTLYLHOYFHI CFVSGMRKIKTAVIGAVYKALVITNSAKRSSTV 420
DB 361 WQGFYFYVLLFVTACLOTLYLHOYFHI CFVSGMRKIKTAVIGAVYKALVITNSAKRSSTV 420
QY 421 GEIYVNLMSVDAQRFMDLATYINMTWSAPLQVILALYLLMLNGPSVLAVAVMTVMEVN 480
DB 421 GEIYVNLMSVDAQRFMDLATYINMTWSAPLQVILALYLLMLNGPSVLAVAVMTVMEVN 480
QY 481 AVMAKTKTYOVAAHKSKDNRIKLMNELNGIKVLKYAMELAFQDKYLAIRQELKYLK 540
DB 481 AVMAKTKTYOVAAHKSKDNRIKLMNELNGIKVLKYAMELAFQDKYLAIRQELKYLK 540
QY 541 KSAVLSAVGTFVWCTPFLVALCTPAVYVTTIDENNIILAOAFVSLAFNIRPPLNTLP 600
DB 541 KSAVLSAVGTFVWCTPFLVALCTPAVYVTTIDENNIILAOAFVSLAFNIRPPLNTLP 600
QY 601 MVISIVQASVSLKRLRIFLSHELEPDSIERRPVKGCGTNSITVRNATFTWASDPPT 660

DB 601 MVISIVQASVSLKRLRIFLSHELEPDSIERRPVKGCGTNSITVRNATFTWASDPPT 660
QY 661 LNCITFSIPREGALVAVVQVCCGSKSLSLALAEEMDKYEGVAIKGSVAAYVPOQAWIOND 720
DB 661 LNCITFSIPREGALVAVVQVCCGSKSLSLALAEEMDKYEGVAIKGSVAAYVPOQAWIOND 720
QY 721 SLRENILFGCQLEPPYRSYIQAACALLPDEILPSGDRTEIGEKVNLSSGQKORVSLAR 780
DB 721 SLRENILFGCQLEPPYRSYIQAACALLPDEILPSGDRTEIGEKVNLSSGQKORVSLAR 780
QY 781 AVYENADITYLEPDDLAVDAVAGCHI PENNYGPKMKIKTRILIVTSMSTLPQVDYIY 840
DB 781 AVYENADITYLEPDDLAVDAVAGCHI PENNYGPKMKIKTRILIVTSMSTLPQVDYIY 840
QY 841 MSGGKISEMSYQELIARDGAFAEFLRTYASTEQODAEENGVTGVSQPGKEALOMENGM 900
DB 841 MSGGKISEMSYQELIARDGAFAEFLRTYASTEQODAEENGVTGVSQPGKEALOMENGM 900
QY 901 LVYDSAGKQLOROLSSSSSYSGDISRHHNSTAEIQAEBAXKEETWKLMEADKAQTGOYKL 960
DB 901 LVYDSAGKQLOROLSSSSSYSGDISRHHNSTAEIQAEBAXKEETWKLMEADKAQTGOYKL 960
QY 961 SVYWDYMKALGLFISFLISIFLMCNHVSALASNYWLSLMTDDPIVNGTQETHKVLSSYG 1020
DB 961 SVYWDYMKALGLFISFLISIFLMCNHVSALASNYWLSLMTDDPIVNGTQETHKVLSSYG 1020
QY 1021 ALGISOGIAVGYMAVSIIGIILASRCIHDVLIHSIRPMSFEPTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVGYMAVSIIGIILASRCIHDVLIHSIRPMSFEPTPSGNLVNRFSEKL 1080
QY 1081 DTVDMSIPEVIMKMGSLFNVIACIYLLATPIAIIIPPLGIYFFVQGFYVASSRQL 1140
DB 1081 DTVDMSIPEVIMKMGSLFNVIACIYLLATPIAIIIPPLGIYFFVQGFYVASSRQL 1140
QY 1141 KRLSEVSRSPPYSHFNETLGVSYIYRAPEQERITHOSDLKVDENOKAYYSIYANRWLA 1200
DB 1141 KRLSEVSRSPPYSHFNETLGVSYIYRAPEQERITHOSDLKVDENOKAYYSIYANRWLA 1200
QY 1201 VRLCEVNCIYLFALFAVVISRHSLSAGLVGLSVSYSLQVTTYINMLVIRMSSEMETINVA 1260
DB 1201 VRLCEVNCIYLFALFAVVISRHSLSAGLVGLSVSYSLQVTTYINMLVIRMSSEMETINVA 1260
QY 1261 VERLKEYSETKEAPWQIQTAPPSWPQVGRVFRNYCLARYRBDLPVLAHINVTINGG 1320
DB 1261 VERLKEYSETKEAPWQIQTAPPSWPQVGRVFRNYCLARYRBDLPVLAHINVTINGG 1320
QY 1321 EKVGI VGRGTGAKSLTILGLFRINESABGEI IIDGINIAKIGLHDLRKTIIIPQDPVLF 1380
DB 1321 EKVGI VGRGTGAKSLTILGLFRINESABGEI IIDGINIAKIGLHDLRKTIIIPQDPVLF 1380
QY 1381 SGLRMLNDPFSQVSDSEVWTSLELAHLKDPVSALPKLDHECAEGGENSVGQROLVCL 1440
DB 1381 SGLRMLNDPFSQVSDSEVWTSLELAHLKDPVSALPKLDHECAEGGENSVGQROLVCL 1440
QY 1441 ARALIRKTKILVDEATAVADLETTDILQSTIRTOFQEDCVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ARALIRKTKILVDEATAVADLETTDILQSTIRTOFQEDCVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIQEYGA PSDLLQORGLFYMAKDAGLV 1531
DB 1501 DKGEIQEYGA PSDLLQORGLFYMAKDAGLV 1531

RESULT 6
US-10-667-891-6
; Sequence 6, Application US/10667891
; Publication No. US20040171024A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BRET, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRAILLES, MARINE


```
; APPLICANT: RZHEITSKY, ANDREY
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
; FILE OF INVENTION: ANOPHELES
; FILE REFERENCE: 03495.0294-00000
; CURRENT APPLICATION NUMBER: US/10/667,891
; CURRENT FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: 60/413,469
; PRIOR FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 6
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-667-891-6

Query Match      78.5%; Score 7860; DB 4; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALRGFCSADGSDPLMDMNTMTNSNDPTGCGONTLVVWPCFYLMACPFYFLYLSRH 60
DB      1 MALRGFCSADGSDPLMDMNTMTNSNDPTGCGONTLVVWPCFYLMACPFYFLYLSRH 60
QY      61 DRGYIQMTPLNKTALGFLMIVCMADLFYSFMEBSRGIFLAVFLVSPFLIGITLLA 120
DB      61 DRGYIQMTPLNKTALGFLMIVCMADLFYSFMEBSRGIFLAVFLVSPFLIGITLLA 120
QY      121 TFLIQERRRGVSGGIMLTFWVALVCAALILRSKIMTALKEDAOVDLFRDITFYFYS 180
DB      121 TFLIQERRRGVSGGIMLTFWVALVCAALILRSKIMTALKEDAOVDLFRDITFYFYS 180
QY      181 LLLIQVLSCGSDSPLEFSETHIDNPNCPRESSASFLSRITFWWTGILVNGRPLEGSD 240
DB      181 LLLIQVLSCGSDSPLEFSETHIDNPNCPRESSASFLSRITFWWTGILVNGRPLEGSD 240
QY      241 LMSLNKEDTSEQVVPVVKWKECAKTRKOPVAVVSSKOPAPKESKVDANEVEAL 300
DB      241 LMSLNKEDTSEQVVPVVKWKECAKTRKOPVAVVSSKOPAPKESKVDANEVEAL 300
QY      301 IVKSPQKEMNPSPLEKVLKTFGPYFLMSFFPKA IHDLMFSGPOLIKLILKFVNDTRAPD 360
DB      301 IVKSPQKEMNPSPLEKVLKTFGPYFLMSFFPKA IHDLMFSGPOLIKLILKFVNDTRAPD 360
QY      361 WQGFYTVLLFVTRACLOTLVHQYFHI CFVSGMRKIAVIGAVYRKALVITNSARKSTV 420
DB      361 WQGFYTVLLFVTRACLOTLVHQYFHI CFVSGMRKIAVIGAVYRKALVITNSARKSTV 420
QY      421 GEIYNLMSVDAOREMDLATYINMWSAPLOVITLALYLLMNLGSPVLAGVAVVWLVMPVN 480
DB      421 GEIYNLMSVDAOREMDLATYINMWSAPLOVITLALYLLMNLGSPVLAGVAVVWLVMPVN 480
QY      481 AVAMAKTKTYQVAHMKSKNRKIKLMEILNGIKVILKYLAMELAFKDVLAIRBELKVLK 540
DB      481 AVAMAKTKTYQVAHMKSKNRKIKLMEILNGIKVILKYLAMELAFKDVLAIRBELKVLK 540
QY      541 KSAVLAVGFTFWCTPFLVALCTFAVYVITDENNNILDAOTAFVSLAFNILREPLNLP 600
DB      541 KSAVLAVGFTFWCTPFLVALCTFAVYVITDENNNILDAOTAFVSLAFNILREPLNLP 600
QY      601 MVTSSIVQASVSLKRLIFLSHELEPDSIERRPVKGKGGNISTVNNATFTWARSPT 660
DB      601 MVTSSIVQASVSLKRLIFLSHELEPDSIERRPVKGKGGNISTVNNATFTWARSPT 660
QY      661 LNGITFESIPGALVAVVGVCGKSSLLSALLAEMDRVESHVAKSGVAVVPOQAMIOND 720
DB      661 LNGITFESIPGALVAVVGVCGKSSLLSALLAEMDRVESHVAKSGVAVVPOQAMIOND 720
QY      721 SLRENIIFGCOLLEPPYRSYIOACALLPDLEILPSGGRTEIGEGVNLSSGQKORVSLAR 780
DB      721 SLRENIIFGCOLLEPPYRSYIOACALLPDLEILPSGGRTEIGEGVNLSSGQKORVSLAR 780
QY      781 AVYSNADIYLFDDPLSAVDHVGHIPEVNIIGPKGMLKNTKTRILVTHSMYSYLPQVDYIY 840
```

```
DB      781 AVYSNADIYLFDDPLSAVDHVGHIPEVNIIGPKGMLKNTKTRILVTHSMYSYLPQVDYIY 840
QY      841 MSGGKISEMKSYOELARDQAFAPLRTYASTEOQDAENGTGVSQPKKAKOMENGM 900
DB      841 MSGGKISEMKSYOELARDQAFAPLRTYASTEOQDAENGTGVSQPKKAKOMENGM 900
QY      901 LVTDSAGKOLQORLSSSSYSYSGDI SRHNSHTAELOKAEKKEFTWKLEADKQOTGVKL 960
DB      901 LVTDSAGKOLQORLSSSSYSYSGDI SRHNSHTAELOKAEKKEFTWKLEADKQOTGVKL 960
QY      961 SYWYDWYKAIGLFISFLIFLMCNHVSALASNYWLSLMTDDPIVNGTOEHTRKLSVYG 1020
DB      961 SYWYDWYKAIGLFISFLIFLMCNHVSALASNYWLSLMTDDPIVNGTOEHTRKLSVYG 1020
QY      1021 ALGISQIAVFGYSMAVSIIGIILASRCLHYDLHSILRSMSFPERTPSGNLVNRSKEL 1080
DB      1021 ALGISQIAVFGYSMAVSIIGIILASRCLHYDLHSILRSMSFPERTPSGNLVNRSKEL 1080
QY      1081 DTWDSMIPVYIKNFMGSLFNVIGACTVILATPIAIIIPPLGIYFVQRFVASSROL 1140
DB      1081 DTWDSMIPVYIKNFMGSLFNVIGACTVILATPIAIIIPPLGIYFVQRFVASSROL 1140
QY      1141 KRLESYSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANWLA 1200
DB      1141 KRLESYSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANWLA 1200
QY      1201 VRLCVCNCIYLPALPAVISRHSLSAGLVGVSYSLOVTTYANLVNRSSEMETNIVA 1260
DB      1201 VRLCVCNCIYLPALPAVISRHSLSAGLVGVSYSLOVTTYANLVNRSSEMETNIVA 1260
QY      1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVPEFRNYCLRYEBDLDFVLRHNVNTNG 1320
DB      1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVPEFRNYCLRYEBDLDFVLRHNVNTNG 1320
QY      1321 EKVIGVGRGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB      1321 EKVIGVGRGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
QY      1381 SSSLRNMLDPSQYSDBEWTSLLELAHLKDFVSALPDKLDHECAGEGENLSVQORQVLCL 1440
DB      1381 SSSLRNMLDPSQYSDBEWTSLLELAHLKDFVSALPDKLDHECAGEGENLSVQORQVLCL 1440
QY      1441 ARALLKRTKTLVDEATAAVDLETTDLOSTIFQEPDCTVLTARLNTIMYTRIVL 1500
DB      1441 ARALLKRTKTLVDEATAAVDLETTDLOSTIFQEPDCTVLTARLNTIMYTRIVL 1500
QY      1501 DKGEIOYGAPSDILOQGLFYSMADQAGLV 1531
DB      1501 DKGEIOYGAPSDILOQGLFYSMADQAGLV 1531

RESULT 7
US-10-807-466-6
; Sequence 6, Application US/10807466
; Publication No. US20040244066A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BREY, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRALLIES, MARINE
; APPLICANT: RZHEITSKY, ANDREY
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
; FILE REFERENCE: 03495.0294-01000
; CURRENT APPLICATION NUMBER: US/10/807,466
; CURRENT FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 10/667,891
; PRIOR FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: 60/413,469
; PRIOR FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn Ver. 3.2
```

SEQ ID NO 6
; LENGTH: 1531
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-807-466-6

Query Match 78.5%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALRGFCADGSDPLMDMNTMTNSNDPTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60
DB 1 MALRGFCADGSDPLMDMNTMTNSNDPTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60
QY 61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPEFLVSPILLGITTTLLA 120
DB 61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPEFLVSPILLGITTTLLA 120
QY 121 TPLIQLEERRKGVOSGIMLTFWLVALVCLALILRSKIMTALKEDAQVDLFRDITFYVYFS 180
DB 121 TPLIQLEERRKGVOSGIMLTFWLVALVCLALILRSKIMTALKEDAQVDLFRDITFYVYFS 180
QY 181 LLLIQVLVSCFSDSPLFSETIHDNPPCESASFLSITTFMWITGLIVRGYROPLEGSD 240
DB 181 LLLIQVLVSCFSDSPLFSETIHDNPPCESASFLSITTFMWITGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKKBCAKTRKOPVKVYSSKDPAPCKESSKYDANEVEVAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKKBCAKTRKOPVKVYSSKDPAPCKESSKYDANEVEVAL 300
QY 301 IVKSPQKEMNPFLKVLKTTGTPYLMSPFFKAIHDLMPFSGPOLIKLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPFLKVLKTTGTPYLMSPFFKAIHDLMPFSGPOLIKLIFVNDTKAPD 360
QY 361 MGGFYTYTLFVYTLCTLOTVLVHOYHIFCVSGMRKKTAVICAVYKAVKAVITNSARKSTV 420
DB 361 MGGFYTYTLFVYTLCTLOTVLVHOYHIFCVSGMRKKTAVICAVYKAVKAVITNSARKSTV 420
QY 421 GEIYVMSVDAQRFMDLATYINMISAPLOVILALYLLMLNGPSVLAVAVMTLVAVN 480
DB 421 GEIYVMSVDAQRFMDLATYINMISAPLOVILALYLLMLNGPSVLAVAVMTLVAVN 480
QY 481 AVMAKTKTYVAVHAKSKDNRIKLMNEILNGIKVILKYAMELAEFKOVLAIROEBLKVLK 540
DB 481 AVMAKTKTYVAVHAKSKDNRIKLMNEILNGIKVILKYAMELAEFKOVLAIROEBLKVLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNIIDAQTAFFVSLAFNLIRPLNLLP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNIIDAQTAFFVSLAFNLIRPLNLLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
QY 661 LMGITFSPICBALVAVVGVQVCGKSSLLSALLAEMDKVEGHVALKGYAVVYPOQAMIGND 720
DB 661 LMGITFSPICBALVAVVGVQVCGKSSLLSALLAEMDKVEGHVALKGYAVVYPOQAMIGND 720
QY 721 SLRENIILFGCOLLEERYRSVIOACALLPDLLEILPSGDTEIGEKVNSISGOKORVSLAR 780
DB 721 SLRENIILFGCOLLEERYRSVIOACALLPDLLEILPSGDTEIGEKVNSISGOKORVSLAR 780
QY 781 AVYNSADIYLFDDPLSAVDANVGHKIFENVIGPKMLKNKTRILVTHGMSYLPQVDVIV 840
DB 781 AVYNSADIYLFDDPLSAVDANVGHKIFENVIGPKMLKNKTRILVTHGMSYLPQVDVIV 840
QY 841 MSGKISMSGSYOELLARDGAFAEFLRTYASTEOBDAEENGVTGVSFGPKAKEMNGM 900
DB 841 MSGKISMSGSYOELLARDGAFAEFLRTYASTEOBDAEENGVTGVSFGPKAKEMNGM 900
QY 901 LVTSAGKQOLOROLSSSSSSSGDISSRHNSTAELQKAKKEETWKLMEADKAOQGVYKL 960
DB 901 LVTSAGKQOLOROLSSSSSSSGDISSRHNSTAELQKAKKEETWKLMEADKAOQGVYKL 960
```

```
QY 961 SVYWDYKKAIGLFIISFLSIFLFMCNHYSAALASNYMLSLMTDDPIVNGTOEHTKVLRSYVG 1020
DB 961 SVYWDYKKAIGLFIISFLSIFLFMCNHYSAALASNYMLSLMTDDPIVNGTOEHTKVLRSYVG 1020
QY 1021 ALGISOGIAVFGYSMAVSGIILASRCILYDILHSILRSMPSPFERIPSGULVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYSMAVSGIILASRCILYDILHSILRSMPSPFERIPSGULVNRFSKEL 1080
QY 1081 DTVDMSIPEVYKMGMSLFNVIGACIVILATPIAIIIPPLGLIYFVORFYAASSROL 1140
DB 1081 DTVDMSIPEVYKMGMSLFNVIGACIVILATPIAIIIPPLGLIYFVORFYAASSROL 1140
QY 1141 KRLESVRSPPYSHFNLTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVRSPPYSHFNLTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLFCVGNCTIVLPALFAVISRHSLSAGLVGLSVSYLOTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLFCVGNCTIVLPALFAVISRHSLSAGLVGLSVSYLOTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPPSMPOVGRVFRNYCLRYREDLPVLRIHINTINGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPPSMPOVGRVFRNYCLRYREDLPVLRIHINTINGG 1320
QY 1321 EKVGIVRGTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRGTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SSGIRPMNLDPPSOYSDEEWTSLELAHLKDFVSLPKLHECAGEGENISVGOROLVCL 1440
DB 1381 SSGIRPMNLDPPSOYSDEEWTSLELAHLKDFVSLPKLHECAGEGENISVGOROLVCL 1440
QY 1441 ABALIRKTKILVDEAPAAVDLEETDDLIOSTIRPOFEDCTVLTAAHLNTIMDTRYIVL 1500
DB 1441 ABALIRKTKILVDEAPAAVDLEETDDLIOSTIRPOFEDCTVLTAAHLNTIMDTRYIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGFLFYSMAKDAGLV 1531
DB 1501 DKGEIOEGAPSDLLQORGFLFYSMAKDAGLV 1531

RESULT 8
US-10-484-577-678
; Sequence 678, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft
; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGT1A1
; FILE REFERENCE: P2285PCT-1
; CURRENT APPLICATION NUMBER: US/10/484,577
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: PCT/EP 02/08220
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: EP 01 11 7608.8
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: EP 02011710.7
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 683
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 678
; LENGTH: 1531
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-484-577-678

Query Match 78.5%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 61 DRGIOMTPLNKTALGFLMIWCMADLFYSFWEBSRGIFLAPVFLVSPILLGITTLA 120
DB 61 DRGIOMTPLNKTALGFLMIWCMADLFYSFWEBSRGIFLAPVFLVSPILLGITTLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMLVVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYS 180
DB 121 TFLIOLERRKGVSSGIMLTFMLVVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYS 180
QY 181 LLLIOLVLSGSDSPLESETIHDPNCPRESSASPLSTITWMTTGLIVRGYRQPLESSD 240
DB 181 LLLIOLVLSGSDSPLESETIHDPNCPRESSASPLSTITWMTTGLIVRGYRQPLESSD 240
QY 241 LMSLNKEDTSBOVAVLVKMKKECAKTRKOPVKVYVSSKOPAPKSSSKVDANEVEAL 300
DB 241 LMSLNKEDTSBOVAVLVKMKKECAKTRKOPVKVYVSSKOPAPKSSSKVDANEVEAL 300
QY 301 IVKSPKEMWNSLFKVLKTPGYFLMSFFPKAHLDMFSGPOLKLLIFVNDTYAPD 360
DB 301 IVKSPKEMWNSLFKVLKTPGYFLMSFFPKAHLDMFSGPOLKLLIFVNDTYAPD 360
QY 361 WOGFYTVLLFVTKLQTLVHQYFHICFVSGKRIKTAIVGAVYRKALVTNSARKSTV 420
DB 361 WOGFYTVLLFVTKLQTLVHQYFHICFVSGKRIKTAIVGAVYRKALVTNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMTMSAPLOVILALYLMLNLSGSLVAGAVMTLMPVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMTMSAPLOVILALYLMLNLSGSLVAGAVMTLMPVN 480
QY 481 AYMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDKVLAIROBELKVLK 540
DB 481 AYMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDKVLAIROBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTPFAVYVITDENNILDQAQFVSLAFNIRPLENLP 600
DB 541 KSAVLSAVGTFTWCTPPLVALCTPFAVYVITDENNILDQAQFVSLAFNIRPLENLP 600
QY 601 MVLSIYQASVSLRILFLSHEELPDSIERPVKGGGNSITVNAATPTMARSPPT 660
DB 601 MVLSIYQASVSLRILFLSHEELPDSIERPVKGGGNSITVNAATPTMARSPPT 660
QY 661 LMGITFSIPRGALVAVVGOCGKSSLSALLAEMDKVGHVAKGSVAVYPOQAMQND 720
DB 661 LMGITFSIPRGALVAVVGOCGKSSLSALLAEMDKVGHVAKGSVAVYPOQAMQND 720
QY 721 SLRENIIIFGCOLEPPYRSVIOACALLPDLLEILPSGRTIEGKGNVLSGGQKQVSLAR 780
DB 721 SLRENIIIFGCOLEPPYRSVIOACALLPDLLEILPSGRTIEGKGNVLSGGQKQVSLAR 780
QY 781 AVYSNADIVLEDDPLASVDAVGHKIFENYIGPKMKLKNTRILVTHSMSTLPQVDYIV 840
DB 781 AVYSNADIVLEDDPLASVDAVGHKIFENYIGPKMKLKNTRILVTHSMSTLPQVDYIV 840
QY 841 MSGGKISEMGSYOELIARDGAFAEFLRTYASTBOODAEENGVTGSGPKKAKOMENG 900
DB 841 MSGGKISEMGSYOELIARDGAFAEFLRTYASTBOODAEENGVTGSGPKKAKOMENG 900
QY 901 LVTDSAGKOLOROLSSSSSYSGDISRHHNSTALLOKAEKKEETWKLMEADKAQTQVXL 960
DB 901 LVTDSAGKOLOROLSSSSSYSGDISRHHNSTALLOKAEKKEETWKLMEADKAQTQVXL 960
QY 961 SYVMDWKAIGLFISEFLIFMGNHVSALASNYLSLMDPDPVNGTOEHTRKLSVYG 1020
DB 961 SYVMDWKAIGLFISEFLIFMGNHVSALASNYLSLMDPDPVNGTOEHTRKLSVYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCIAYDLHSILRSPMSFEERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCIAYDLHSILRSPMSFEERTPSGNLVNRFSEKL 1080
QY 1081 DYTDSMTPEYIKPMGSLFNVIKACIYIILATPIAIIIPPLGLIYFVQRFYVASSROL 1140
DB 1081 DYTDSMTPEYIKPMGSLFNVIKACIYIILATPIAIIIPPLGLIYFVQRFYVASSROL 1140
QY 1081 DYTDSMTPEYIKPMGSLFNVIKACIYIILATPIAIIIPPLGLIYFVQRFYVASSROL 1140

QY 1141 KRLSEVRSPPVSHFNELLGVSVIRAFEEQEREIHOSDKVDENOKAYPPIVANRWLA 1200
DB 1141 KRLSEVRSPPVSHFNELLGVSVIRAFEEQEREIHOSDKVDENOKAYPPIVANRWLA 1200
QY 1201 VRLCEVNCIYLFALFAVISRHSLSAGVLVSYSLOVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLFALFAVISRHSLSAGVLVSYSLOVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEMPMQIOETAPPSWPOVGRVFRNYCLARREDLDFVLRHINVTNG 1320
DB 1261 VERLKEYSETEKEMPMQIOETAPPSWPOVGRVFRNYCLARREDLDFVLRHINVTNG 1320
QY 1321 EKVIGVGTAGKSSLTGLFRINESAGEIIGINIAKIGLHDLRFKTIIPQDPYLF 1380
DB 1321 EKVIGVGTAGKSSLTGLFRINESAGEIIGINIAKIGLHDLRFKTIIPQDPYLF 1380
QY 1381 SGLRNMIDPPSOYSDSEWTSLELAHLKDFVSALPDKLHCEKAGEENLSVGROLVCL 1440
DB 1381 SGLRNMIDPPSOYSDSEWTSLELAHLKDFVSALPDKLHCEKAGEENLSVGROLVCL 1440
QY 1441 ARALRRTKILVDEATAVADLETDLIQSTIRQPEDCTVLTARHANTIMDYTRYVL 1500
DB 1441 ARALRRTKILVDEATAVADLETDLIQSTIRQPEDCTVLTARHANTIMDYTRYVL 1500
QY 1501 DKGEIOEGAPSDILOQRGLFYSMAKXAGLV 1531
DB 1501 DKGEIOEGAPSDILOQRGLFYSMAKXAGLV 1531

RESULT 9
US-10-889-503-19
Sequence 19, Application US/10889503
Publication No. US20050063968A1
GENERAL INFORMATION:
APPLICANT: Fox Chase Cancer Center
APPLICANT: Krush, Gary D.
APPLICANT: Lee, Kun
APPLICANT: Belinsky, Martin G.
TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
FILE REFERENCE: FCCC 98-02
CURRENT APPLICATION NUMBER: US/10/889,503
PRIOR APPLICATION NUMBER: US/09/647,140
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US99/06644
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/079,759
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/095,153
PRIOR FILING DATE: 1998-08-03
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-889-503-19
Query Match 78.5%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGCSADGSDPLMDMNTVNTSNPDPFKCFQNTVLVWPCFYLMACPPFFLYSRH 60
DB 1 MALRGCSADGSDPLMDMNTVNTSNPDPFKCFQNTVLVWPCFYLMACPPFFLYSRH 60
QY 61 DRGIOMTPLNKTALGFLMIWCMADLFYSFWEBSRGIFLAPVFLVSPILLGITTLA 120
DB 61 DRGIOMTPLNKTALGFLMIWCMADLFYSFWEBSRGIFLAPVFLVSPILLGITTLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMLVVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYS 180

```
|||||
Db 121 TFLQLERKRGVSSGIMLTFMLVALYCALAILRSKIMTALKEDAQVDLFPFDITFYVVS 180
Qy 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRTFMWITGLIVRGROPLEGSD 240
Db 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRTFMWITGLIVRGROPLEGSD 240
Qy 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPKSSSKVDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPKSSSKVDANEVEAL 300
Qy 301 IVKSPQKEMNSLPKVLKTFGPFYFLMSFPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNSLPKVLKTFGPFYFLMSFPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Qy 361 WQGFYTVLFLVTAQLOTVLVHQYFHIQFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYTVLFLVTAQLOTVLVHQYFHIQFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
Qy 421 GEIVNLSVDAQRFPMDLATYINMISAPLOVILALYLLMLNLSGVSILAGVAVMVLMEVN 480
Db 421 GEIVNLSVDAQRFPMDLATYINMISAPLOVILALYLLMLNLSGVSILAGVAVMVLMEVN 480
Qy 481 AVMAAMKTTYQVAAHKKSKONRIKLMNEITLNGIKVILKYAMELAFKQVLAIRQBELKVLK 540
Db 481 AVMAAMKTTYQVAAHKKSKONRIKLMNEITLNGIKVILKYAMELAFKQVLAIRQBELKVLK 540
Qy 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNILLAQAFVSLFNLIRPPLNLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNILLAQAFVSLFNLIRPPLNLP 600
Qy 601 MVISIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGGNTSITVRNATFTWASDPT 660
Db 601 MVISIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGGNTSITVRNATFTWASDPT 660
Qy 661 LMGITFSLPEGALVAVVQVCGGKSSLLSALLAENDKVBGHAIGSAVYVPOQAMTND 720
Db 661 LMGITFSLPEGALVAVVQVCGGKSSLLSALLAENDKVBGHAIGSAVYVPOQAMTND 720
Qy 721 SLRENILGCGOLEBEYYRSVIOACALLPDLLETLPBGDTEIGEKKVNLSSGGKQKVSILAR 780
Db 721 SLRENILGCGOLEBEYYRSVIOACALLPDLLETLPBGDTEIGEKKVNLSSGGKQKVSILAR 780
Qy 781 AVYSNADITLFPDDLPSAVDAHVGKHI FENVIGPKMKLNKTRILVTHSMSYLPQVDIIV 840
Db 781 AVYSNADITLFPDDLPSAVDAHVGKHI FENVIGPKMKLNKTRILVTHSMSYLPQVDIIV 840
Qy 841 MSGKISMGSYOELLARDGAFAEFLRTYASTEBOQDAEENGVTGVSQKKEAKOMENGM 900
Db 841 MSGKISMGSYOELLARDGAFAEFLRTYASTEBOQDAEENGVTGVSQKKEAKOMENGM 900
Qy 901 LVTHSAGOLOROSSSSSGDTSRHHNSTAELOKAKKKEETKLMBADDAQOQVUL 960
Db 901 LVTHSAGOLOROSSSSSGDTSRHHNSTAELOKAKKKEETKLMBADDAQOQVUL 960
Qy 961 SVYMDYMAKIGLIFSLIFLFCMCHVSLASNYMLSLMTDPIVNGTOEHTKVLSTYVG 1020
Db 961 SVYMDYMAKIGLIFSLIFLFCMCHVSLASNYMLSLMTDPIVNGTOEHTKVLSTYVG 1020
Qy 1021 ALGISQGIIVFGYSMAVSIIGIILASRCLHVDLHLSILKSPMSFPERTPSGNLVNFSKEL 1080
Db 1021 ALGISQGIIVFGYSMAVSIIGIILASRCLHVDLHLSILKSPMSFPERTPSGNLVNFSKEL 1080
Qy 1081 DTVMSIMEVIMKMGSLFENVIGACIVILLATFPIAIIIPGLIYFPVQRYVASSQOL 1140
Db 1081 DTVMSIMEVIMKMGSLFENVIGACIVILLATFPIAIIIPGLIYFPVQRYVASSQOL 1140
Qy 1141 KRLSEVSRSPVYSHNETLIGSVVIRAFEBEORFIHQSDLKVDENOKAYPSIVANRMLA 1200
Db 1141 KRLSEVSRSPVYSHNETLIGSVVIRAFEBEORFIHQSDLKVDENOKAYPSIVANRMLA 1200
Qy 1201 VRLBCVNCIVLFAALFAVISHSHSISAGLVGSVSYSLQVTTYINMLVMSSEMETNIVA 1260
Db 1201 VRLBCVNCIVLFAALFAVISHSHSISAGLVGSVSYSLQVTTYINMLVMSSEMETNIVA 1260
```

```
Db 1201 VRLBCVNCIVLFAALFAVISHSHSISAGLVGSVSYSLQVTTYINMLVMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFRNYCLRYREDDLVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFRNYCLRYREDDLVLRHINVTINGG 1320
Qy 1321 ERKGIIVGRTGAGKSLTLGLFRINESAGEIINDGINIAKIGLHDLRPKITIIIPQDPVLF 1380
Db 1321 ERKGIIVGRTGAGKSLTLGLFRINESAGEIINDGINIAKIGLHDLRPKITIIIPQDPVLF 1380
Qy 1381 SGSLRMLNDPSSQVSDDEWMTSLAHLKQPVSAIPKTLHECAGEGENLSVGOQOLVCL 1440
Db 1381 SGSLRMLNDPSSQVSDDEWMTSLAHLKQPVSAIPKTLHECAGEGENLSVGOQOLVCL 1440
Qy 1441 ARALLRKTILVDEATAVADLETDDLIOSTIRTOFEDCVLITAHRLNTIMDYTRVVL 1500
Db 1441 ARALLRKTILVDEATAVADLETDDLIOSTIRTOFEDCVLITAHRLNTIMDYTRVVL 1500
Qy 1501 DKGEIOEYGAPSDILLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDILLQORGLFYSMAKDAGLV 1531

RESULT 10
US-10-756-149-5033
; Sequence 5033, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Zichnik, Natasha
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5033
; LENGTH: 1531
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-756-149-5033

Query Match 78.5%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCGADGSDPLMDMNTWNTSNPDFTKCFONTVLVWVPCRYLMACFPFYFLYLSRH 60
Db 1 MALRGFCGADGSDPLMDMNTWNTSNPDFTKCFONTVLVWVPCRYLMACFPFYFLYLSRH 60
Qy 61 DRGYIOMPLNKKTALGFLIMIVCMADLFYSFMERSGIFLAVPLVSPULLGITTLA 120
Db 61 DRGYIOMPLNKKTALGFLIMIVCMADLFYSFMERSGIFLAVPLVSPULLGITTLA 120
Qy 121 TFLQLERKRGVSSGIMLTFMLVALYCALAILRSKIMTALKEDAQVDLFPFDITFYVVS 180
Db 121 TFLQLERKRGVSSGIMLTFMLVALYCALAILRSKIMTALKEDAQVDLFPFDITFYVVS 180
Qy 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRTFMWITGLIVRGROPLEGSD 240
Db 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRTFMWITGLIVRGROPLEGSD 240
Qy 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPKSSSKVDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPKSSSKVDANEVEAL 300
Qy 301 IVKSPQKEMNSLPKVLKTFGPFYFLMSFPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNSLPKVLKTFGPFYFLMSFPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Qy 361 WQGFYTVLFLVTAQLOTVLVHQYFHIQFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYTVLFLVTAQLOTVLVHQYFHIQFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
```

Db 361 WQGYFYTVLLFVTAQLOTLVLHQYFHCFSVGNRIKTAIVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAORFMDLATYTNMTWSAPLOYTLALYLLMNLGSSVLAGVAVMLAMPVN 480
Db 421 GEIYNLMSVDAORFMDLATYTNMTWSAPLOYTLALYLLMNLGSSVLAGVAVMLAMPVN 480
Qy 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAKDKVLAIROBELKVLK 540
Db 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAKDKVLAIROBELKVLK 540
Qy 541 KSAVLASVGTFTWCTPPLVALCTFPAYVTIDENNILDAQTAFAVSLAFNIRPPLNLP 600
Db 541 KSAVLASVGTFTWCTPPLVALCTFPAYVTIDENNILDAQTAFAVSLAFNIRPPLNLP 600
Qy 601 MVTSSIVQASVSLKRLIFLSHELEPDSIFERRPVKGGGNSITVNNATFTWARSDDPT 660
Db 601 MVTSSIVQASVSLKRLIFLSHELEPDSIFERRPVKGGGNSITVNNATFTWARSDDPT 660
Qy 661 LNCITFPIPEGALVAVGVCGSGSSLSALLAEMDKVGHVAKGSVAVYPOQAMQND 720
Db 661 LNCITFPIPEGALVAVGVCGSGSSLSALLAEMDKVGHVAKGSVAVYPOQAMQND 720
Qy 721 SLRENILFGCQLEBPYRSVIAQCALLPDLLEILPSGDRTEIGEKVNLGGQKQVSLAR 780
Db 721 SLRENILFGCQLEBPYRSVIAQCALLPDLLEILPSGDRTEIGEKVNLGGQKQVSLAR 780
Qy 781 AVYSNADIYLPDPLSAVDHVGKHTFENVIGPFGMLKNTKRIIVTHSMSTLPQVDYIIV 840
Db 781 AVYSNADIYLPDPLSAVDHVGKHTFENVIGPFGMLKNTKRIIVTHSMSTLPQVDYIIV 840
Qy 841 MSGGKISEMGSYOBLARDGAPAEFLRTYASTBEOADABENGVTGSGPGKEAKOMENG 900
Db 841 MSGGKISEMGSYOBLARDGAPAEFLRTYASTBEOADABENGVTGSGPGKEAKOMENG 900
Qy 901 LVTDLSACKOLORLSSSSSSSGDISRHHNSTABLOKAEKKEETWKMMEADKQGTQVKL 960
Db 901 LVTDLSACKOLORLSSSSSSSGDISRHHNSTABLOKAEKKEETWKMMEADKQGTQVKL 960
Qy 961 SVYWDYKAIKGLFISLIFLPMGNHVSALASNYWLSMTDDPIVNGTOEHTKRLSVYG 1020
Db 961 SVYWDYKAIKGLFISLIFLPMGNHVSALASNYWLSMTDDPIVNGTOEHTKRLSVYG 1020
Qy 1021 ALGISOGIAVFGVMAVSIIGLILASRCLHYDLHSLIRSPMSFEERPPSGNLVNRFSKEL 1080
Db 1021 ALGISOGIAVFGVMAVSIIGLILASRCLHYDLHSLIRSPMSFEERPPSGNLVNRFSKEL 1080
Qy 1081 DTVDMSIPEYIKHMGSLFNVIKACIYILATPIAIIIPPLGLIYFVORFYVASSROL 1140
Db 1081 DTVDMSIPEYIKHMGSLFNVIKACIYILATPIAIIIPPLGLIYFVORFYVASSROL 1140
Qy 1141 KRLSVSRSPVYSHFNFTLLGVSYIRAPBEOERFIHQSDLKVDENQAYYPSIVANWLA 1200
Db 1141 KRLSVSRSPVYSHFNFTLLGVSYIRAPBEOERFIHQSDLKVDENQAYYPSIVANWLA 1200
Qy 1201 VRLBECVNCI.V.PAALPAVISRHSLSAGLVGSYSLOYTYIANMIVRMSSEMETNIVA 1260
Db 1201 VRLBECVNCI.V.PAALPAVISRHSLSAGLVGSYSLOYTYIANMIVRMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKAPWQIOETAPPSWPQVGRVFRNYCLARYEDDLFVLRHINVTINGG 1320
Db 1261 VERLKEVSETEKAPWQIOETAPPSWPQVGRVFRNYCLARYEDDLFVLRHINVTINGG 1320
Qy 1321 EKVGIVERTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDVLV 1380
Db 1321 EKVGIVERTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDVLV 1380
Qy 1381 SSGLRNMLDPPSOYSDEEWTSLLEAHLKDFVSLPKLDBHECGEGENISVGRQOLVCL 1440
Db 1381 SSGLRNMLDPPSOYSDEEWTSLLEAHLKDFVSLPKLDBHECGEGENISVGRQOLVCL 1440
Qy 1441 ARALLRKTKLIVLDEATAVADLETDLLIOSTIRTOFEDCTVLTAHSLANTIMDYTRYIVL 1500
Db 1441 ARALLRKTKLIVLDEATAVADLETDLLIOSTIRTOFEDCTVLTAHSLANTIMDYTRYIVL 1500

Qy 1501 DKGEIOBYGAPSDLLQORGLFYSGAKDAGLV 1531
Db 1501 DKGEIOBYGAPSDLLQORGLFYSGAKDAGLV 1531

RESULT 11
US-09-939-853A-87
; Sequence 87, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT FILING DATE: US/09/939, 853A
; PRIOR APPLICATION NUMBER: 2001-08-27
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ. ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-87

Query Match 77.6%; Score 7769; DB 3; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 DMVNTNTSPDPTKCGQNTVWVWPCFYIMACFPFYLYLSRHDGYIOMTPLNKTCTA 76
Db 1 DMVNTNTSPDPTKCGQNTVWVWPCFYIMACFPFYLYLSRHDGYIOMTPLNKTCTA 60
Qy 77 LGFLMIVCADLIFYFWERSRGIPLAPVLSPTLLGTTLATFLIQLERRKGVSSG 136
Db 61 LGFLMIVCADLIFYFWERSRGIPLAPVLSPTLLGTTLATFLIQLERRKGVSSG 120
Qy 137 IMLPMLVALVCAIIRSKIMTALKEDAVDLFRDITFYVYFSLLIQVLSCFSRSP 196
Db 121 IMLPMLVALVCAIIRSKIMTALKEDAVDLFRDITFYVYFSLLIQVLSCFSRSP 180
Qy 197 LPSETIHDPRCPRESSASFLSRITFWITGLYRGYRQPLEGSDLSINKEDTSEOVVPV 256
Db 181 LPSETIHDPRCPRESSASFLSRITFWITGLYRGYRQPLEGSDLSINKEDTSEOVVPV 240
Qy 257 LVNWMKECKAKTKQPKVYVSSKDPAPQPKSSKVDANEVEALIVKSPCKENNPISLFKV 316
Db 241 LVNWMKECKAKTKQPKVYVSSKDPAPQPKSSKVDANEVEALIVKSPCKENNPISLFKV 300
Qy 317 LYTFPGPYFLMSPFFKAIDHLMFSGPOIILKLIKFNVDTKADWQGYFTVLLFVYACL 376
Db 301 LYTFPGPYFLMSPFFKAIDHLMFSGPOIILKLIKFNVDTKADWQGYFTVLLFVYACL 360
Qy 377 QTLVHLQYFHICVSGMRITKTAIVGAVYRKALVITNSARSSIVGEIVNLMSVDAORFMD 436
Db 361 QTLVHLQYFHICVSGMRITKTAIVGAVYRKALVITNSARSSIVGEIVNLMSVDAORFMD 420
Qy 437 LATYINMISAPQVILALYILMLNGPSVLAGVAVVWLVVPAVWAMKTKTYOVAMHK 496
Db 421 LATYINMISAPQVILALYILMLNGPSVLAGVAVVWLVVPAVWAMKTKTYOVAMHK 480
Qy 497 SKDNRIKLMNEILNGIKVLKLYAMELAKDKVLAIROBELKVLKKSAYVSAVGTFTWCT 556
Db 481 SKDNRIKLMNEILNGIKVLKLYAMELAKDKVLAIROBELKVLKKSAYVSAVGTFTWCT 540
Qy 557 PFLVALCTFPAVYVTIDENNILDAQTAFAVSLAFNIRPPLNIPMTVSSIVQASVSLKRL 616

```
|||||
Db      541 PFLVALCTFAVYVTTIDENNIIIDAQTAFAVSLAFNILRPEPLILPMVISSIVQASVSKRL 600
Qy      617 RIFLSHELEPDSIERRPVVDGGGNSITVNAATFTWARSPTPLNGTTSIPGALVAV 676
Db      601 RIFLSHELEPDSIERRPVVDGGGNSITVNAATFTWARSPTPLNGTTSIPGALVAV 660
Qy      677 VGVGCGKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGQLEBPY 736
Db      661 VGVGCGKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGQLEBPY 720
Qy      737 YRSVIOACALLPDLIELPSGDRTEIGEGVNLSGGQKRVSLARAVYSNADIIYFDDPLS 796
Db      721 YRSVIOACALLPDLIELPSGDRTEIGEGVNLSGGQKRVSLARAVYSNADIIYFDDPLS 780
Qy      797 AYDAHVGHIPENYIGPKMKKNKTRILVTHSMSTLPQVDYIIVMSGKISEMSYOBLL 856
Db      781 AYDAHVGHIPENYIGPKMKKNKTRILVTHSMSTLPQVDYIIVMSGKISEMSYOBLL 840
Qy      857 ARDGAFAFLRTVASTEOBDAEENGVTGVSQPKBAKOMENGMVYDTSAGKOLQROLSS 916
Db      841 ARDGAFAFLRTVASTEOBDAEENGVTGVSQPKBAKOMENGMVYDTSAGKOLQROLSS 900
Qy      917 SSSYSGDISRHNSTAELOKAEAKKEETWKLMEADKAOTGOYKLSVYWDYMKALGLFTSF 976
Db      901 SSSYSGDISRHNSTAELOKAEAKKEETWKLMEADKAOTGOYKLSVYWDYMKALGLFTSF 960
Qy      977 LSIPLFMCNHYSAASNYWMLSLMTDDPIVNGTQOETHKRLSVYGALGISOGIAVFGSMA 1036
Db      961 LSIPLFMCNHYSAASNYWMLSLMTDDPIVNGTQOETHKRLSVYGALGISOGIAVFGSMA 1020
Qy      1037 VSIGIILASRCLAHVDLHSLRSPPSPERTPSGNLVNRFSKELDTVDMSIPEVIKPMWG 1096
Db      1021 VSIGIILASRCLAHVDLHSLRSPPSPERTPSGNLVNRFSKELDTVDMSIPEVIKPMWG 1080
Qy      1097 SLENVIGACTIYLLATPIAAIIIPPLGIYFFVQAFYASSRQLKRLSVSRSPYSHFN 1156
Db      1081 SLENVIGACTIYLLATPIAAIIIPPLGIYFFVQAFYASSRQLKRLSVSRSPYSHFN 1140
Qy      1157 ETLIGVSVIRAFEOERFIHOSDLKVDENOKAYYPSIYANRMLVRLCVCNCIYLPAL 1216
Db      1141 ETLIGVSVIRAFEOERFIHOSDLKVDENOKAYYPSIYANRMLVRLCVCNCIYLPAL 1200
Qy      1217 FAVISRHSLSAGVLGVSYSLOVTTYLNMLVRMSSEMETNIIVAVERLKEYSETEKEAPW 1276
Db      1201 FAVISRHSLSAGVLGVSYSLOVTTYLNMLVRMSSEMETNIIVAVERLKEYSETEKEAPW 1260
Qy      1277 QIOETAPSSWPQVGRVFEFRNYCLRYREDLDFVLPHINVTINGGEKVGIVERTGAGKSSL 1336
Db      1261 QIOETAPSSWPQVGRVFEFRNYCLRYREDLDFVLPHINVTINGGEKVGIVERTGAGKSSL 1320
Qy      1337 TLGIFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLPESGSLRMLDPPSOYSD 1396
Db      1321 TLGIFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLPESGSLRMLDPPSOYSD 1380
Qy      1397 EEVWTSLELAHKDFVSALPKLDHECABEGENISVQOROLVCLARALLRKTILVILDEA 1456
Db      1381 EEVWTSLELAHKDFVSALPKLDHECABEGENISVQOROLVCLARALLRKTILVILDEA 1440
Qy      1457 TAAVDELTDULIOSTIRTOFEDCTYLLTAHRLNTIMDYTRYIVLIDKGEIOEYGAESDILQ 1516
Db      1441 TAAVDELTDULIOSTIRTOFEDCTYLLTAHRLNTIMDYTRYIVLIDKGEIOEYGAESDILQ 1500
Qy      1517 ORGLFYSAKADAGLV 1531
Db      1501 ORGLFYSAKADAGLV 1515
```

```
APPLICANT: Day, Anthony G.
APPLICANT: Estell, David A.
APPLICANT: Lyons, Eric H.
APPLICANT: Yao, Jian
TITLE OR INVENTION: Methods for Modulating Proteins Not
TITLE OR INVENTION: Previously Known as Proteases
FILE REFERENCE: GC773-2
CURRENT APPLICATION NUMBER: US/10/618,281
PRIOR APPLICATION DATE: 2003-07-11
PRIOR APPLICATION NUMBER: US 60/395,325
NUMBER OF SEQ ID NOS: 92
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 42
LENGTH: 1515
TYPE: PR
ORGANISM: Homo sapiens
US-10-618-281-42
```

```
Query Match 77.5%; Score 7763; DB 5; Length 1515;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      17 DNANTNTSNPDFTKCHQNTVLYVWPCFYLMACPFYFLYSRDRGTYQMTPLNKTITA 76
Db      1 DNANTNTSNPDFTKCHQNTVLYVWPCFYLMACPFYFLYSRDRGTYQMTPLNKTITA 60
Qy      77 LGFLLMIVCANADLFYSWERSRGIFLAPVFAVSPTLLGITTLATPFIQLERRGVQSSG 136
Db      61 LGFLLMIVCANADLFYSWERSRGIFLAPVFAVSPTLLGITTLATPFIQLERRGVQSSG 120
Qy      137 IMLFMLVALVICALAILRSKIMTALKEDAQVDLFRDITFYVYFSLILQVLSCFSRSP 196
Db      121 IMLFMLVALVICALAILRSKIMTALKEDAQVDLFRDITFYVYFSLILQVLSCFSRSP 180
Qy      197 LFSSTIHDPNCPBSSASFSRITFMWITGLIYRGYQPLBGSJLMSLNKEDTSEQVVPV 256
Db      181 LFSSTIHDPNCPBSSASFSRITFMWITGLIYRGYQPLBGSJLMSLNKEDTSEQVVPV 240
Qy      257 LVNKKWKECAKTRQAPKVVVYSSKDPAPKSSKVDNNEFEALIVSPKEMNPSLFKV 316
Db      241 LVNKKWKECAKTRQAPKVVVYSSKDPAPKSSKVDNNEFEALIVSPKEMNPSLFKV 300
Qy      317 LYKTFEYFLMSFFFKAIHDLMFSGPQILKLIKFNVDTKAPDMQGYFTVLLFVTAQL 376
Db      301 LYKTFEYFLMSFFFKAIHDLMFSGPQILKLIKFNVDTKAPDMQGYFTVLLFVTAQL 360
Qy      377 QTVLHGYFHI CFVSGNRITKTAIVGAYYRKALVITNSARKSSTVGEIVNLSVDAQRMD 436
Db      361 QTVLHGYFHI CFVSGNRITKTAIVGAYYRKALVITNSARKSSTVGEIVNLSVDAQRMD 420
Qy      437 LATYNNINWASAPLOVILALVILMLNLPVSLAGAVWVWLVVNAWAMTKTKYQVAHMK 496
Db      421 LATYNNINWASAPLOVILALVILMLNLPVSLAGAVWVWLVVNAWAMTKTKYQVAHMK 480
Qy      497 SKDNRIKLMEIILNGIKVLKYAMELAFKQVLAIROEBELVKLKSAYLSAVGTFTWCT 556
Db      481 SKDNRIKLMEIILNGIKVLKYAMELAFKQVLAIROEBELVKLKSAYLSAVGTFTWCT 540
Qy      557 PFLVALCTFAVYVTTIDENNIIIDAQTAFAVSLAFNILRPEPLILPMVISSIVQASVSKRL 616
Db      541 PFLVALCTFAVYVTTIDENNIIIDAQTAFAVSLAFNILRPEPLILPMVISSIVQASVSKRL 600
Qy      617 RIFLSHELEPDSIERRPVVDGGGNSITVNAATFTWARSPTPLNGTTSIPGALVAV 676
Db      601 RIFLSHELEPDSIERRPVVDGGGNSITVNAATFTWARSPTPLNGTTSIPGALVAV 660
Qy      677 VGVGCGKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGQLEBPY 736
Db      661 VGVGCGKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGQLEBPY 720
Qy      737 YRSVIOACALLPDLIELPSGDRTEIGEGVNLSGGQKRVSLARAVYSNADIIYFDDPLS 796
```



```

Db      721 YRSVIOACALLPDLLEILPSGDRTEIGKGNVLSGGQKQVSLAARAVSNADLYLFDPLLS 780
Qy      797 AVDAHVGKHEFENYTGPKGMLKNKTRILVTHSMSTLPQVDVITVWSGKISEMSYQELL 856
Db      781 AVDAHVGKHEFENYTGPKGMLKNKTRILVTHSMSTLPQVDVITVWSGKISEMSYQELL 840
Qy      857 ARDGAFAEFLRTYASTEOBDAENGVTGVSQPKGAQKQEMNGMLVTDSACKOLQORLSS 916
Db      841 ARDGAFAEFLRTYASTEOBDAENGVTGVSQPKGAQKQEMNGMLVTDSACKOLQORLSS 900
Qy      917 SSSYSGDISRRHNSYAELOKAEKKEETWKLMEADKAQTGVQKLSVVDYMKALGLFISF 976
Db      901 SSSYSGDISRRHNSYAELOKAEKKEETWKLMEADKAQTGVQKLSVVDYMKALGLFISF 960
Qy      977 LSLFLPMCNHVSALASNTWLSMTDDPIVNGTOEHTKRLSVYGAAGISQGIANFGISMA 1036
Db      961 LSLFLPMCNHVSALASNTWLSMTDDPIVNGTOEHTKRLSVYGAAGISQGIANFGISMA 1020
Qy      1037 VSIGGILASRCLAVDLHSILRSPMSFEPTSPGNLVNRFSKBLDTVDSMTPEVTKMFMG 1096
Db      1021 VSIGGILASRCLAVDLHSILRSPMSFEPTSPGNLVNRFSKBLDTVDSMTPEVTKMFMG 1080
Qy      1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQKRLSESVRSPPYSHFN 1156
Db      1081 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQKRLSESVRSPPYSHFN 1140
Qy      1157 ETLIGSVYIRAFEEQERFIHOSDLKVDENQKAYIPSIYANFVLARLECGVNCIVLFAAL 1216
Db      1141 ETLIGSVYIRAFEEQERFIHOSDLKVDENQKAYIPSIYANFVLARLECGVNCIVLFAAL 1200
Qy      1217 PAVISRSLSAGVIGLSVSYSLQVTTYTNMLVRMSSEMETIVAVERKESSETEKAPW 1276
Db      1201 PAVISRSLSAGVIGLSVSYSLQVTTYTNMLVRMSSEMETIVAVERKESSETEKAPW 1260
Qy      1277 QIOETAPSSWPQYGRVEFRNYCLRYEDDLFVLRHINVTINGEKYQIVGRTAGKSSL 1336
Db      1261 QIOETAPSSWPQYGRVEFRNYCLRYEDDLFVLRHINVTINGEKYQIVGRTAGKSSL 1320
Qy      1337 TLGLFRINESAGEHIIIDGINIAKIGLHDLRFKTIIPQDPVLFGSLRMNLDPFSQYSD 1396
Db      1321 TLGLFRINESAGEHIIIDGINIAKIGLHDLRFKTIIPQDPVLFGSLRMNLDPFSQYSD 1380
Qy      1397 BEWVTSLELAHLKQFVSALPKLDEHCEGEGENISVQOROUVCARALARTKILVNDDEA 1456
Db      1381 BEWVTSLELAHLKQFVSALPKLDEHCEGEGENISVQOROUVCARALARTKILVNDDEA 1440
Qy      1457 TAAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVLDKGEIOEYGAPSDLLQ 1516
Db      1441 TAAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVLDKGEIOEYGAPSDLLQ 1500
Qy      1517 QRGFLFYMAKDAGLV 1531
Db      1501 QRGFLFYMAKDAGLV 1515

```

```

RESULT 13
US-10-408-765A-1718
; Sequence 1718, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fany, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408.765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1718
; LENGTH: 1388
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-408-765A-1718

Query Match      70.3%; Score 7037; DB 4; Length 1388;
Beet Local Similarity 97.0%; Pred. No. 0;
Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;

Qy      67 MTPANKTKTALGFLMTIVCMADLYFSFMEBSRGIFLAPVFLVSPTLGITTLLATPLIQ 126
Db      1 MTPANKTKTALGFLMTIVCMADLYFSFMEBSRGIFLAPVFLVSPTLGITTLLATPLIQ 60
Qy      127 ERRGVSSGIMLTFWMLVALCALAIRSKIMTAKEDAQVLDPRDITFYVYFSLILQL 186
Db      61 ERRGVSSGIMLTFWMLVALCALAIRSKIMTAKEDAQVLDPRDITFYVYFSLILQL 97
Qy      187 VLSCFSRSLPFSETHIDPNPCPSSASFLSRITFMWITGLIYNGYQPLEGSDLMSLNK 246
Db      98 -----NCPSSASFLSRITFMWITGLIYNGYQPLEGSDLMSLNK 138
Qy      247 EDTSEQVVPVLVKNWKKCAKTRQPKVYSSKDPAPQPKSSKVDANBEVEBALIVKSPQ 306
Db      139 EDTSEQVVPVLVKNWKKCAKTRQPKVYSSKDPAPQPKSSKVDANBEVEBALIVKSPQ 198
Qy      307 KENNPSLFKVLKTFGYPFLMSPEFKAIDHLMFSGPOLIKLIFVNDTKAPDMQGYFY 366
Db      199 KENNPSLFKVLKTFGYPFLMSPEFKAIDHLMFSGPOLIKLIFVNDTKAPDMQGYFY 258
Qy      367 TVLLFTVACIQTLYLHOYFHI CFVSGKRIKTAIVGAYRKALVITNSARKSSTYGEIYNL 426
Db      259 TVLLFTVACIQTLYLHOYFHI CFVSGKRIKTAIVGAYRKALVITNSARKSSTYGEIYNL 318
Qy      427 MSYDAORFMDLATYINMWSAPLOVILALYLMNTGPSYLAGAVANVLVNVANWAMK 486
Db      319 MSYDAORFMDLATYINMWSAPLOVILALYLMNTGPSYLAGAVANVLVNVANWAMK 378
Qy      487 TKTYQVAMKSKONRIKLANEIIINGIKVLKLYAMELAFKDKVLAIROBELKVLKKSAYLS 546
Db      379 TKTYQVAMKSKONRIKLANEIIINGIKVLKLYAMELAFKDKVLAIROBELKVLKKSAYLS 438
Qy      547 AVGTFTVCTPPLVALCTFAVYVTTIDENNILDAQTAFVSLAFNIIIRPLNIIIPMVISI 606
Db      439 AVGTFTVCTPPLVALCTFAVYVTTIDENNILDAQTAFVSLAFNIIIRPLNIIIPMVISI 498
Qy      607 VQASVSLKRLRIFLSHEELEPPDSIERRPVMDGGTNSITVRNATFTWABSDPPTLNGITF 666
Db      499 VQASVSLKRLRIFLSHEELEPPDSIERRPVMDGGTNSITVRNATFTWABSDPPTLNGITF 558
Qy      667 SIEGALVAVAVGVGCGKSSLSALLAEMDKVEGHVAKISGVAVVPOQAMTQNDSLRENI 726
Db      559 SIEGALVAVAVGVGCGKSSLSALLAEMDKVEGHVAKISGVAVVPOQAMTQNDSLRENI 618
Qy      727 LFGQULEEPPYRSYVQACALLPDLLEILPSGDRTEIGKGNVLSGGQKQVSLAARAVY SNA 786
Db      619 LFGQULEEPPYRSYVQACALLPDLLEILPSGDRTEIGKGNVLSGGQKQVSLAARAVY SNA 678
Qy      787 DYLFDPLSLAVDAHVGKHEFENYTGPKGMLKNKTRILVTHSMSTLPQVDVITVWSGKI 846
Db      679 DYLFDPLSLAVDAHVGKHEFENYTGPKGMLKNKTRILVTHSMSTLPQVDVITVWSGKI 738
Qy      847 SEMGSYQELLARDAFAEFLRTYASTEOBDAENGVTGVSQPKGAQKQEMNGMLVTDSACK 906
Db      739 SEMGSYQELLARDAFAEFLRTYASTEOBDAENGVTGVSQPKGAQKQEMNGMLVTDSACK 798
Qy      907 GKOLQORLSSSSSGDISRRHNSYAELOKAEKKEETWKLMEADKAQTGVQKLSVVDY 966
Db      799 GKOLQORLSSSSSGDISRRHNSYAELOKAEKKEETWKLMEADKAQTGVQKLSVVDY 858
Qy      967 MKAIGFISFLSIFLPMCNHVSALASNTWLSMTDDPIVNGTOEHTKRLSVYGAAGISQ 1026

```

Db 859 MKAIGLPISTFLFMCNHNVSALASNYWLSLMTDDPIVNGTQEHTRKLSVYGALGISQ 918
Qy 1027 GIAVFGYSMAVSIIGIILASRCILAVDLHLSILRSPMSFFERTPSGNLVNRFSEKELDTVDSM 1086
Db 919 GIAVFGYSMAVSIIGIILASRCILAVDLHLSILRSPMSFFERTPSGNLVNRFSEKELDTVDSM 978
Qy 1087 IPEYIKMFMGSLFNVIAGACIYILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSEV 1146
Db 979 IPEYIKMFMGSLFNVIAGACIYILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSEV 1038
Qy 1147 SRSVYSHNFETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYYSIVANMFLAVRECV 1206
Db 1039 SRSVYSHNFETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYYSIVANMFLAVRECV 1098
Qy 1207 GNCIVLFAALFAVISRHSLSAGLVGLSVSYSLOVTTYLNMLVRMSEMETNIIVAVERLKE 1266
Db 1099 GNCIVLFAALFAVISRHSLSAGLVGLSVSYSLOVTTYLNMLVRMSEMETNIIVAVERLKE 1158
Qy 1267 YSETEKEAPMOIOETAPSSNPQVGRVFRNYCLARYEDLDPLRHINVTINGEKVGI 1326
Db 1159 YSETEKEAPMOIOETAPSSNPQVGRVFRNYCLARYEDLDPLRHINVTINGEKVGI 1218
Qy 1327 GRTAGKSLTLGLFRINESABGRIIDGINIAKIGLHDLRPKTIIIPDPLVPSGSLRM 1386
Db 1219 GRTAGKSLTLGLFRINESABGRIIDGINIAKIGLHDLRPKTIIIPDPLVPSGSLRM 1278
Qy 1387 NLDPFSQYSDEEVTSTELAHKDFVSALPKDLHCEKAGENLSVQGRQVLCLARALLR 1446
Db 1279 NLDPFSQYSDEEVTSTELAHKDFVSALPKDLHCEKAGENLSVQGRQVLCLARALLR 1338
Qy 1447 KTKILVLDEATPAVDLETTDLIOSTIRIQFEDCTVLTIAHRLNTIMDYTR 1496
Db 1339 KTKILVLDEATPAVDLETTDLIOSTIRIQFEDCTVLTIAHRLNTIMDYTR 1388

RESULT 14

US-09-939-853A-88
; Sequence 88, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OR INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 1528
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-88

Query Match 69.9%; Score 7002.5; DB 3; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

Qy 1 MALGFGSADGSDPLMDNNTWNTSNEDPTKCFONTVLVWPCEFYLAACPFYFYLSRH 60
Db 1 MALGFGSADGSDPLMDNNTWNTSNEDPTKCFONTVLVWPCEFYLAACPFYFYLSRH 60
Qy 61 DRGIOMPLUNKTKALGFLLMIWCMADLFYSFMRSHRGIFLAPVFLVSPPTLIGITTTLLA 120
Db 61 DRGIOMTHLUNKTKALGFLLMIWCMADLFYSFMRSHRGIFLAPVFLVSPPTLIGITTTLLA 120

Qy 121 TFLIQLERRKGVOSSGIMLTFWLVALVICALAIILRSKTMALKEDAQVDLFRDITFPYFYS 180
Db 121 TFLIQLERRKGVOSSGIMLTFWLVALVICALAIILRSKTMALKEDAQVDLFRDITFPYFYS 180
Qy 181 LLLIQVLSCGSDSPILFSETIHDNCPRESSASFRLTFPMWITGLIIVRGQRPLGSSD 240
Db 181 LVLVQLVLSGSDSPILFSETIHDNCPRESSASFRLTFPMWITGLIIVRGQRPLGSSD 240
Qy 241 LMSLNKEDTSEQVVPVLVWKMKCECAKTRKQPVNVYVS-SKDPAPQKSSKVDANEVEYA 299
Db 241 LMSLNKEDTSEQVVPVLVWKMKCECAKTRKQPVNVYVS-SKDPAPQKSSKVDANEVEYA 300
Qy 300 LIVSPQKMNPSLFXVLKTFPGPYFLMSFFPKAIDMLMPSGPQILKLIFVNDITKAP 359
Db 301 LIVSPHKBDEBPSLFXVLKTFPGPYFLMSFFPKAIDMLMPSGPQILKLIFVNDITKAP 360
Qy 360 DMQGFYFTVLLFVYACIOTLVHGYFHLCPVSGRITAVYGAVYRALKVITNSARSSST 419
Db 361 DMQGFYFTVLLFVYACIOTLVHGYFHLCPVSGRITAVYGAVYRALKVITNSARSSST 420
Qy 420 VGEIYVNLMSYDAQRFMDLATYINMISAPQVILATYLMNLGSPVLAGVAVVLMVVP 479
Db 421 VGEIYVNLMSYDAQRFMDLATYINMISAPQVILATYLMNLGSPVLAGVAVVLMVVP 480
Qy 480 NAVAMKTKTYQVAHMSKONRIKLAMBIINGIVLKLAMELAPKOKVLAIROBELKVL 539
Db 481 NAVAMKTKTYQVAHMSKONRIKLAMBIINGIVLKLAMELAPKOKVLAIROBELKVL 540
Qy 540 KKSAYVLAVGFETVWCPELVALCTPFAVYVITDENNTILDQTAVSLAFPIITL 599
Db 541 KKSAYVLAVGFETVWCPELVALCTPFAVYVITDENNTILDQTAVSLAFPIITL 600
Qy 600 PMVSSIVQASVSLKRLRIFLSHELEPDSIERPVRDGGGTNSITVYNAFTWARSDDP 659
Db 601 PMVSSIVQASVSLKRLRIFLSHELEPDSIERPVRDGGGTNSITVYNAFTWARSDDP 659
Qy 660 TLNGITTSIPBGALVAVGVGCGKSSLSALAEKMDKVBGHVAKSVAVYPOQAWTON 719
Db 660 TLNGITTSIPBGALVAVGVGCGKSSLSALAEKMDKVBGHVAKSVAVYPOQAWTON 719
Qy 720 DSLAEKNTLFGCOUEEPYRYSYQACALLPDLIELPDSDRTEIGKGVNLSCGQKORVSLA 779
Db 720 DSLAEKNTLFGCOUEEPYRYSYQACALLPDLIELPDSDRTEIGKGVNLSCGQKORVSLA 779
Qy 780 RAYVSNADIYLPDPLSAVDAAVGHKIFENVIYCGKMLKNKTRILVYTHSMYSILPOVDYII 839
Db 780 RAYVSNADIYLPDPLSAVDAAVGHKIFENVIYCGKMLKNKTRILVYTHSMYSILPOVDYII 839
Qy 840 VMSGGKISEMGSYQELLARDGAPAEFLRTYASTQEOBDAEENGVTGVSGPKKAKQWENG 899
Db 840 VMSGGKISEMGSYQELLARDGAPAEFLRTYASTQEOBDAEENGVTGVSGPKKAKQWENG 896
Qy 900 MLVYDSAGKQLOROLSSSSSYSGDISRHHNSTALQRAAKKETWLMEDAKQOTGVYK 959
Db 900 MLVYDSAGKQLOROLSSSSSYSGDISRHHNSTALQRAAKKETWLMEDAKQOTGVYK 955
Qy 897 MLVYDVGKHLQRRHLSNSSHSGDTSOQHSIAELQAGA-KEBTWKLMEDAKQOTGVYK 955
Qy 960 LSYVMDVWKAIGLISFLSTFLFMCNHNVSALASNYWLSLMTDD-PYNGTQEHTRKRLSV 1018
Db 956 LSYVMDVWKAIGLISFLSTFLFMCNHNVSALASNYWLSLMTDDPPVNGTQAHKRFRLSV 1015
Qy 1019 YGALGISQGIADVFGYSMAVSIIGIILASRCILAVDLHLSILRSPMSFFERTPSGNLVNRSK 1078
Db 1016 YGALGISQGIADVFGYSMAVSIIGIILASRCILAVDLHLSILRSPMSFFERTPSGNLVNRSK 1075
Qy 1079 ELDTVDKMIPEVYIKMFMGSLFNVIAGACIYILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1138
Db 1076 ELDTVDKMIPEVYIKMFMGSLFNVIAGACIYILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1135
Qy 1139 QKRLSESVSRPVYSHNFETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYYSIVANRW 1198
Db 1136 QKRLSESVSRPVYSHNFETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYYSIVANRW 1195
Qy 1199 LAVRLCEVGNCIIVLFAALFAVISRHSLSAGLVGLSVSYSLOVTTYLNMLVRMSEMETNI 1258

Db 1196 LAVERECVNCIVLFAALPAVISRHSLSAGLVGSVSLQITAYLMLVLRMSSEMETNI 1255
Qy 1259 VAVERLKEYSETEAEAWOIOETPSSWPOVGEVERRYNCLARYEDLPYLRHINTIN 1318
Db 1256 VAVERLKEYSETEAEAWOIOETPSSWPOVGEVERRYNCLARYEDLPYLRHINTIN 1315
Qy 1319 GGEKVGIVGRTRGAKSSLTGLFRLNSAEGEIITDGINIAKIGLHDLRFKITTIPDDV 1378
Db 1316 GGEKVGIVGRTRGAKSSLTGLFRLNSAEGEIITDGINIAKIGLHDLRFKITTIPDDV 1375
Qy 1379 LFGSGLRMLNDPFSQYSDERWTSLELAHLNDPVSALPDKLDHBCAEGENTSVGOROLV 1438
Db 1376 LFGSGLRMLNDPFSQYSDERWTSLELAHLNDPVSALPDKLDHBCAEGENTSVGOROLV 1435
Qy 1439 CLARALRKTKILVLDATAVLDLETDDLIOSTIRTOPEDCTVTIARHLNTIMDYTRVI 1498
Db 1436 CLARALRKTKILVLDATAVLDLETDDLIOSTIRTOPEDCTVTIARHLNTIMDYTRVI 1495
Qy 1499 VLDKGEIOEXGAPSDLIQORGLFYSMAKDGLV 1531
Db 1496 VLDKGEIOEXGAPSDLIQORGLFYSMAKDGLV 1528

RESULT 15
US-10-618-281-63
; Sequence 63, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 1303
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-618-281-63

Query Match 60.6%; Score 6068.5; DB 5; Length 1303;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1229; Conservative 17; Mismatches 57; Indels 207; Gaps 10;

Qy 22 WNTSNPDFTKCFONTVLMVWPCFYLMACPFYFLYSRHDGYIOMTPLNKTKTALGFL 81
Db 1 WNTSNPDFTKCFONTVLMVWPCFYLMACPFYFLYSRHDGYIOMTPLNKTKTALGFL 60

Qy 82 WIYCMADLFYSFERNRSGITLAPVFLVSPILLGITLLATPLIQLERRKVOSSGIMLT 141
Db 61 WIYCMADLFYSFERNRSGITLAPVFLVSPILLGITLLATPLIQLERRKVOSSGIMLT 120

Qy 142 WUVALCALALNSKIMTAKEDAOVDFRDIFFYVYFSLITQVLVSCSDSPLESET 201
Db 121 WUVALCALALNSKIMTAKEDAOVDFRDIFFYVYFSLITQVLVSCSDSPLESET 177

Qy 202 IHDNPNCPESSASFLSRITFMWITGLIVRGYROPLESGDLSINKEDTSEQVVPVLYKN 261
Db 178 IHDNPNCPESSASFLSRITFMWITGLIVRGYROPLESGDLSINKEDTSEQVVPVLYKN 237

Qy 262 KKECAKTRKQPVKVVYVSSKDPADPKESSKYDANEVEVALVKSPOKEMNPSLFKULYKT 321
Db 238 KKECAKTRN-----SSGSGESCSANTALF-----PA--PTCHKSF 271

Qy 322 GPYFLMSFFPKALHDLMMFSGPOLIKLLIFVNDTKAPDWQGYFYTVLLEVTACLQTLV 381

Db 272 QALSL-----LCRLIKFVNDTKAPDWQGYFYTVLLEVTACLQTLV 314
Qy 382 HOYHICFVSGMKIKTAVICAYRKALVTNSARKSTVGEIYNLMSVDAQREMDLATYI 441
Db 315 HOYHICFVSGMKIKTAVICAYRKALVTNSARKSTVGEIYNLMSVDAQREMDLATYI 374
Qy 442 NMMSAPLOVILALYLMLNLGSPVLAVGVVWMLMPVNVVAMKTKTYVAHKKSDNR 501
Db 375 NMMSAPLOVILALYLMLNLGSPVLAVGVVWMLMPVNVVAMKTKTYVAHKKSDNR 416
Qy 502 IKLMEINLNGIKVLKYAMELAFKDYLAIRQELKYLKKSAYLSAVGTFTWCTPPLVA 561
Db 417 IKLMEINLNGIKVLKYAMELAFKDYLAIRQELKYLKKSAYLSAVGTFTWCTPPLVA 476
Qy 562 LCTPAVYVTTIDENNILDAQTAFLVSLALFNLRPLNLPVVISIYQASVSLKRLIFLS 621
Db 477 LCTPAVYVTTIDENNILDAQTAFLVSLALFNLRPLNLPVVISIYQASVSLKRLIFLS 528
Qy 622 HEELEPDSIERPVKQCGGTSITVNATFTWASDPPLNGITFSPREGALVAVGVONG 681
Db 529 GATSEKGPWGSRRKRG-----TRQASFSVABPGVLCRFSTFSPREGALVAVGVONG 581
Qy 682 CGKSSLLSALLAEMDKVEGVAIKGSVAVYPOQAMIONDSLRENILFGCOLLEPYRSVI 741
Db 582 CGKSSLLSALLAEMDKVEGVAIKGSVAVYPOQAMIONDSLRENILFGCOLLEPYRSVI 641
Qy 742 QACALLPDLLEILPSGDRTEIGERGVNLGGQOKRVSLARAVSNADITLFDPLSAVDH 801
Db 642 QACALLPDLLEILPSGDRTEIGERGVNLGGQOKRVSLARAVSNADITLFDPLSAVDH 701
Qy 802 VGHIFPENYVGPQMKLKNKTRILVYTHSMSTLPQVDVYIYWSGKISMGSTQELIARDGA 861
Db 702 VGHIFPENYVGPQMKLKNKSCLT----- 724
Qy 862 FAEFLPTVASTBQEDAEENGVTGVSQPKAKOMENGMVTDQAGKQLQRLSSSSYS 921
Db 725 -----S 725
Qy 922 GDISRHNSTYALQKAKAKKERTWKLMEADKQOTGVKLSVYDMYKAIGLFISLSIFL 981
Db 726 CDL-----QYKLSVYDMYKAIGLFISLSIFL 753
Qy 982 FMCNHYSAASNYWLSMTDDPLVNGTOHTKRLSYGALGISOGIAVNGYSMAVSG 1041
Db 754 FMCNHYSAASNYWLSMTDDPLVNGTOHTKRLSYGALGISOGIAVNGYSMAVSG 813
Qy 1042 ILASRCIADVLLHSIIRSPMSFFERTPSGNLVNRFSEKLDTVDSMTPEVTKMFGSLFNV 1101
Db 814 ILASRCIADVLLHSIIRSPMSFFERTPSGNLVNRFSEKLDTVDSMTPEVTKMFGSLFNV 873
Qy 1102 IGACIVILLATPIAIIIPPLGLIYFVQRFYVASSROLKRLSESVRSPLYSHNETLIG 1161
Db 874 IGACIVILLATPIAIIIPPLGLIYFVQRFYVASSROLKRLSESVRSPLYSHNETLIG 933
Qy 1162 VSVYIRAFEBQERFIHOSDKVDENOKAYYPSIYANWLAVERLECVGNCTYLPALPAVIS 1221
Db 934 VSVYIRAFEBQERFIHOSDKVDENOKAYYPSIYANWLAVERLECVGNCTYLPALPAVIS 933
Qy 1222 RHLASGLVGLSVSYSLQVTTYVNLVVRMSSEMETNI VAVERLKEYSETEAEAWOIOET 1281
Db 994 RHLASGLVGLSVSYSLQVTTYVNLVVRMSSEMETNI VAVERLKEYSETEAEAWOIOET 1053
Qy 1282 APPSSWPOVGRVFRYNYCLARYEDLPVLRHINVTINGEKVGI VGRTRGAKSSLTGLF 1341
Db 1054 APPSSWPOVGRVFRYNYCLARYEDLPVLRHINVTINGEKVGI VGRTRGAKSSLTGLF 1113
Qy 1342 RINESAEGEITIDGINIAKIGLHDLRFKITTIPDDVLPFSGLRMLNDPFSQYSDERWT 1401
Db 1114 RINESAEGEITIDGINIAKIGLHDLRFKITTIPDDVLPFSGLRMLNDPFSQYSDERWT 1173
Qy 1402 SLELAHKPFVSLPDKLDHBCAEGENTSVGOROLVCLARALRKTKILVLDATAV 1461

Db	1174	SLELAHLKDFVSALPDKLDHECAREGENTSVQOROLVCLAPALAKTKILVLDENTAAND	1233
Qy	1462	LETDDLQSTIRTFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQORGLF	1521
Db	1234	LETDDLQSTIRTFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQORGLF	1293
Qy	1522	YSMAKDAGLV	1531
Db	1294	YSMAKDAGLV	1303

Search completed: December 15, 2005, 15:47:17
Job time : 170.491 secs

Oy	495	MEKORIKAMEINIINGIVLKYAMELAFKOKVIAIQOEELKVLKSAVYSAVGTFFW	55
Db	1	MROKDSRAPUTSSILNSKTIIFKHEMGEAFIDRVGIRHGEELGALRTSGLLFSJSVYFQ	60
Oy	555	CTFFVLALCTFAVYVTTIDENNILDAQAFVSLAFNLLRPLNLLPMVYSSIVQASVSLK	614
Db	61	VSTFELVALVFAVHTLVAE-NAMNAEKAFFVTLVTIININKAAQAFPEFISHLVQARVAFD	114
Oy	615	RLTIFPSHEELBPDSTERRPVKDDGNTSITVRNATFTWASDDPTLNGIRFSPREGALV	674
Db	120	RLVTFELCEBVDGVNDSSSSGAAGKOCITTIQSTTFPMOSEBPCLHRIULVTPQGLL	178
Oy	675	AVVGQVCGCKSSILSLALAEKMDVEGHAIVKGSVAVYVPOQAMIQNDSLRENILFGCOLE	734
Db	180	AVVGPAGAGKSSILSLALGELSKVGFAPISIBGAVAVYVOEAMVQNTSVENAVCGOEIDP	234
Oy	735	PYRSVYIOACALLPDLIELPSGDRIEISEKGNVLSGCGOKORVSLARAVYSAADYTFDDP	794
Db	240	PMELRVLEACALQDPVDSPFEGIIHNSIIEGQGNMSSCGQKQSLSLARAVYRAAVYLLDDP	299
Oy	795	LSAVDAHNGKHIFENVIGPKMKLKKKTKTLLVTHMSNYLPQVDVLLVWMSGKISMSGYOE	854
Db	300	LAALDAHVQOHFNQVITGPGSLIGQFTRTLVTHALHTLLPQADMTIIVLANGIAIEMSGYOE	354

QY 855 LLAARDGAFAEFLRYASTGEODAEENGWTVSGPGKAEKMEMGMLVYDSAGK--OQOR 912
 DB 360 LLOKRGKALVCLL-----DQAROPGDRG--EGETEPGISTKXDPG-----TSAGRRPELRR 407
 QY 913 QLSSSSYSGDISHNHNSSTAELOKAEAKKEET-----KCLMEADKAKOTGOVLSYVWDMK 968
 DB 408 ERSLKS-----VPEKDRITSEAO--TEVPLDDPPDAKMPAK--DSIQGRVATVHALAYLR 460
 QY 969 AIGLIFSLIFLMCNHVSALASNYMLSLMTDDPIVNGTOEHKTVLSVYGALISQGI 1028
 DB 461 AVGPRLCYALFLPLCOQVNASFCRGWMLSLWADDPAGVGQOTQALNRGIGLIGCAI 520
 QY 1029 AVRGYSNAVSIGTILASRCHVDLHLSLSPSGPFERTPSGNTVNRPSKELDVDSMP 1068
 DB 521 GLFASMAAVLILGGARASRLFLQRLMDVRSPISEFFERTPIGHLNRPFSKETDVVDIIP 580
 QY 1089 EVIMFMSLBNVAGACIVILATPIAIIIPGLIYFFQRFVYVASSRDLKLESVR 1148
 DB 581 DKLSLIMYAFGLAEVSLVAVATPLATVALPLFLYAGFQSLYVSSCOLRLESASY 640
 QY 1149 SPVYSHNETLLGVSVIRAFEEORFIHQSDLKVDENOKAYPSIVANRMLAVLECVN 1208
 DB 641 SSVCSHMAETFGSTTVRFRTOAPFVAQNNARVDESQRISFPRILVADRWLAANVELLN 700
 QY 1209 CTVLPALLFVISHHSISAGLVLGSVSYSLQVTTYTLAMLVMSSEMETNIYAVELKEYS 1268
 DB 701 GLVPAATCAVLSKXHLISAGLVGSVSAALQVOTQLQVWVBNMTDLENSIYSVERMODYA 760
 QY 1269 ETEBEAQMIOETPSPSPMPQVGRVPEPNYCLARYBEDDFVLRHNTVINGEKVGIYGR 1328
 DB 761 WTPKEAEMRLPTCAQPPWPQGGQIEFDFGLRYRPELPLAVQGVSEFIHAGEKGIYGR 820
 QY 1329 TGAKSSLLTGLFRINSEAGEIIIDGINIAKIGLHDLRFKITIIPQDPVLFSSSLRML 1388
 DB 821 TGAKSSLASLALHQAEBGCIWIDGVPILHVGHLTRISISIIPODPIFPSSIRML 880
 QY 1389 DPFQSYDEEYVTSLELAHLKDFVSALEPKDLHCEAGSENLVSGOROLVCLARLAKT 1448
 DB 881 DLLOEHSDEALWALETVQLKALVASLFGQLQYKCADGEBLSVGQQLCLARALAKT 940
 QY 1449 KIIVLDEATAVDLETDLOSTIRTOPEDCVTULIARLNTIMDYRIVYLDGEIOEX 1508
 DB 941 QILILDEATAVDGTLEQWQAMLSWPACTVLIARLRSVMDCAVAVLMDKQVAVS 1000
 QY 1509 GAPSDDLQORGLFYSMADAGLV 1531
 DB 1001 GSPAQLAQKGLFYRLAQESGLV 1023

RESULT 2

US-11-090-439-24
 ; Sequence 24, Application US/11090439
 ; Publication No. US20050266442A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Squillace, Rachel
 ; APPLICANT: Weiner, Michael P.
 ; TITLE OF INVENTION: Immortalized Human Tuberculous Sclerosis Null
 ; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
 ; FILE REFERENCE: 24318-502
 ; CURRENT APPLICATION NUMBER: US/11/090,439
 ; PRIOR FILING DATE: 2005-03-25
 ; PRIOR APPLICATION NUMBER: 60/556,344
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 24
 ; LENGTH: 1581
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-090-439-24

Query Match 21.1%; Score 2115; DB 7; Length 1581;
 Best Local Similarity 31.9%; Pred. No. 1,4e-146;

Matches 527; Conservative 305; Mismatches 564; Indels 256; Gaps 37;
 QY 32 CFQNTVLVWVPCFYLMACPFEPFL-YLSRHDRCYIQWTPINKTKTALGF---LMIWCV 86
 DB 26 CFVDALVWVPHVFLFTFPILFIMGSSQSKVHHNS-----TWLHPCHNLRIWLITLF 79
 QY 87 ADLFYSFWESRGIFFLAPV-----FLVSP---LIGITTLATFLIQLERRGVOSGC 136
 DB 80 MLFLVLCESIAEGILSDGVESHHLIYMPAGNAFMAVLSV--YYHNIETSNPKLLI 137
 QY 137 IMLEFWLVALVCMALIRSKIMTALXEDAQVDLFRDITF-----VYFSLILQVLVS 189
 DB 138 ALIVYMTLAFI-----TKIKFVKFLDIAIGFSQRLCTGLLVLYLGMKLVE--VN 188
 QY 190 CFSDRSLFSETHIDPNCDES-----SASFLSRITFWMTGLVYRGROPLEG 238
 DB 189 VIRRRRIYIFETREVPAP--PEDLODLQVRLQPPVNLISKTYMMWNAFIKAAKPI-- 245
 QY 239 SDMLSLNKEDTSBOVVLVK--NMKKECAKTRKQPVKVVYSKDPADQPKESKVDANE 295
 DB 246 -DLRAIK-----LPIAMRALTNYORLC-----E 268
 QY 296 EVELIYKSPQ--KEMNPSLFLVLYKTEGPFYFMSFFPKAIDHLMWFSGPQILKLIKRV- 353
 DB 269 AFDQVQRKDIOGTQARAIWQALSHAFGRRLVLSSTRILADLGFAGPLCIFGIVDHIG 328
 QY 354 --NDTKAPDWQ-----GYFYVLLFVTAQIQTVLVHQYFHCIFVSGRIK 396
 DB 329 KENDVPQPKQFLGVFVVSQEFPLANAYVLAVLLFLLLRTPFLQASYYVAIEGTINLR 388
 QY 397 TAVIGAYYRKA--LVITNSARKSSTGEIYNLMSVDQRFMDLATYINMISAPLOYILA 454
 DB 389 GAIQTKLYNKMIMHLSLNSMGWETAQICNLVAIDNTQMLPFPLCPNLAMAMVOQIIVG 448
 QY 455 LYILMLNGSVLAGVAVMLVMPVNAVMAKKTQYQVAHKSNDRIKLMNEITLNGKV 514
 DB 449 VILLYVILGVSAALGAIVILILAPVOYVANKLSQAPRSLEYSENERIKQTNEMLRGKL 508
 QY 515 LKLYAMELAFQDKVLAIRQEBLKVLKKSAYLSAVGTFWVCTPPLVALCTPAVVYTTDEN 574
 DB 509 LKLYAMENIRTRRETRRREKEMTSIRAFIYTSISIMNTRAPIPAVALITPVGHSFPKE 568
 QY 575 NIIDAQAFVSLAFNLIRPLNLIPVYISIVQASVSLKRLIFLSHELEPDSI--- 630
 DB 569 ADFSPVAFASLSLPHILVLPFLSSVSVSTVALVSVOGLSSEFLSABIREBQCAPHE 628
 QY 631 -----ERRPVK-----DGGTN-SITVANNATPTW 653
 DB 629 PTPQPAKTYQAVPLRVNRRPARBEDRGLTGFLQSLVPSADGDADNCCVQIMGYFT 688
 QY 654 ARSDPTLNGITFSIPGALVAVVGVGCGKSSLSALLAEMDKVEG----- 700
 DB 689 TPDSIPTLSNTIIPRGQLTMIVGVGCGKSSLSLLALGEMQVSGAVFWSLSLPDEIG 748
 QY 701 -----HVAIKSVAVVPPQAMIQNDSLRNLLFGCLSEEPYKSVIQACALL 747
 DB 749 EDPSPERETATDLDIRKGPVAYVQSPMLNATVNEIIPESPFRKORYGVATACSLQ 808
 QY 748 PDLEILSGDRTEIGEGGNVLSGQKORVSLARAVNSADLYLPDDPLSANDAVKGIIF 807
 DB 809 PDIDILPHGOTOIGERGINLSGQORISVARRLYOHANVAVLPDDPSALDHLSDHLM 868
 QY 808 ENVIGPKGMLK-----NKTRLIVTHSASYLPQOVVIVYMSGKISEMSGYOELARDAF 862
 DB 869 Q-----AGILELADDDKGTIVLVTHKQYLPHADWITIAMKDGITQREBTLKDPQSEQL 923
 QY 863 AEFRLTYASTEQODAEENGTVSGPGKEAKQENGLVYDSAGQLQORQLSSSSSYSG 922
 DB 924 FEHWKTLMN-RQDELKEKETY-----ERKATEPPQG-----LSRAMSRDGL-- 965
 QY 923 DISRHNHSTAELOKAEAKKEETWKLMEADKA-----QTGVKLSVYMDYMAIGLIFSL 977
 DB 966 -----LQDEEBEERBAESEBDMNLSSMLHORAEIPWRACAKYLSAGIILLSTL 1014


```

Oy 978 SIPLMCHVSAALSNWMLSMITDDPIV-----NCTOHT-----KYRLSYGAG 1023
Db 1015 LVFSOLKHMVLVADIDYWLAKTDBALTTLPARNCSJSECTLDQIVYAVAFVLSGL 1074
Oy 1024 ISOGIAVEGYSMAVSIIGIILASRCLHVLHLSHSPMSFFERTPSGNLVNRFSEKLDTV 1083
Db 1075 I---VLGVTSVTVWEWTGLKVAKALHRSLSLNRILLAPRFFETTPGLSILNRFSDCMTI 1111
Oy 1084 DSMIPVYIKMFMSGLPNVIGACIVLLATPIAAITIPPLGLITFFVQRFYVASSRQLKRL 1143
Db 1132 DQHPSTLECTSRSLTLCVSAALAVSYTVPFVLALEPLAIVCYFIQKXPFVASRDLQOL 1191
Oy 1144 ESVSRSPYXSHNETILGVSIVARBEEOERFHHOSDKVDENOKXYPSIVANRMLAVRL 1203
Db 1192 DDTTOLPLSHFAETVEBGLTTIRAFRYEARFOQLLETTDSNNIASLFLTAANNVLEVRM 1251
Oy 1204 ECVGNCIVLFAALFAVIS--RHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVAV 1261
Db 1252 EYIGACVVLIAAVTSISNSLHRELSAGLVGGLTVLALVNSYILMMVRNLDLMELQGVAV 1311
Oy 1262 ERLKEYSTEXE-----APMOIQETAPSPPOVGRVFPFNYYCLIRREDDFLYLRHNV 1315
Db 1312 KRIRGLATKEESYEGGLAPSLI-----PKWPPDQKIQIONLSTRYDSSLKPYLKHVNA 1366
Oy 1316 TINGEKYGIWGRTAGKSSLTILGLFRINESABEIIIDGINIAKTGLHDLRPFITTIPO 1375
Db 1367 LISPOKIGICGRYSGKSSSLAFRRVNDTRBHHIIDDIDIAKPLHTTRSRSLILO 1426
Oy 1376 DPVLFSGSLRNLDLBFQYSDEEVTSLLEHLNDQFVASLPDKLDHECAEGENLSVGOR 1435
Db 1427 DPVLFSGTIRFNLDPERKCSSTLMEALEIQKLKVLVALLPGGDAIITEGGENFSGOR 1486
Oy 1436 QLVCLARALKKTKLVLDEHTAANDLETDLDIGSTIRQEDCTVLTIRARLNTINDYT 1495
Db 1487 QLPCLARAFVKRTSLFFIMDEATASIDMATEMILQCVWTAADRVITVIAHVTILISAD 1546
Oy 1496 RVIYLDKGEIOEYGAPSDILOOR--GLPFSMAK 1526
Db 1547 LVIVYKRAILIEFDPKEKTLRSKRQSVFASFVR 1578

RESULT 3
US-10-995-561-527
; Sequence 527, Application US/10995561
; Publication No. US20050272054v1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CU001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 527
; LENGTH: 1588
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-527

```

	Query Match	21.0%	Score 2101.5	DB 6	Length 1568
	Best Local Similarity	31.8%	Pred. No. 14e-145		
	Matches	527	Conservative 305	Mismatches 564	Indels 263 Gaps 38
QY	32	CEQNVLVWPCFYLMACPFYFL	YLSDHGDGYLMTPLNKTALGF	---	LLMTVCM 86
Db	26	CFVDMLNVVPHVFLFTTPTL	FLTGSSOSKVIHHS	----	TWLPHPGNLRILTF 79
QY	87	ADLPYSFWERSKGIPLAVP	----	FLVSPT	---LLGITLLATFLQLERRKGVSSG 136
Db	80	MLFLVAVCEIAGLISDVGTESSHHLLVMPDAGMAFAAATSVV	----	YVNIIEFSNPKPLI	137

QY	137	MLFMFWALVACALALRSKTMTLKEDAOVDLPEDIF-----VYFSLLLQVLVS	189
Db	138	ALVYWTALF-----TKTJKVKFLDHAIGFSQLEFCLTGLLVILYGMILLVE--VN	188
QY	190	CFSRSPFLFSTIHDNPNCPE\$-----SASFSLRTFTWVTGLVIRGYNQPLEG	238
Db	189	VIRVRRYIFFTPREVKP-PEDLODLGVRFLOFPNLLSKGTYYMMNMFITKAHKPI--	245
QY	239	SDLSLNKEDTSEOVNPLVK--NMKECAKTRKQPVKVYSSKDPAPQESSKYDANE	295
Db	246	-DLBAIGK-----LPIAMRALTNYSOLC-----E	268
QY	296	EVEALIVYSPQ-KEMNPFLFKVLKTFEPYFLMSEFFKAIHDLMMFSGPOLKLLIKFY-	353
Db	269	AFDQVNRDIOCTOGARAIMQALSHARBRVLVISTFPIIADLIGFADPLCFGLVDHIG	328
QY	354	--NDTKAPDWQ-----GYFTVLLFVTACIOTVLVHQYFHI CFYSGMRK	396
Db	329	KENDVPQCKTQFLGVYFVSSQEPFLANAVALVLLFLLALLORTFLOASYYVAIEGINLR	388
QY	397	TAVIGAVYRKA--LVITNSARKSSTVGEIVNLM\$VDQRPMDLATYIMIMASBPQVILA	454
Db	389	GAIOTKYTNKTMHLSSTNSLWGEKMTAGQICLVVAIDNOLMMFFFLCPNLNAMPQOITVG	448
QY	455	LYLLMLNGPSVLGAVAMVLMVPMVNAAMAKTQYQVAHAKSKDNRIKLMNLIINGIKV	514
Db	449	VILLYYILGV\$ALIGAAVITILAPVQYVVAKLSQAQ\$STLEYSNERKQJNEMRGIKL	508
QY	515	LKLYAMELAFQDKLAIROEBLKVLK\$SAYISAVGTFTWVCTPPLVALCTPAVYTTIDEN	574
Db	509	LKLYAMENIIFPTRVETTRRKEMTSLRAFAITYSISIFNNTAIPLAAVILITFVGH\$PFKE	568
QY	575	NIILDAQFAV\$ALFENTILRPPLNTLPMWISSIVQASV\$KRLRFL\$HEELEE\$PDI----	630
Db	569	AD\$P\$VA\$FASL\$PHILVTLFLLSSVYR\$T\$KALVSQ\$KLS\$EFL\$SABE\$REBCA\$HE	628
QY	631	-----ERRPVK-----DGGGTN-SITVRNATFTW	653
Db	629	PTQGPASKYQAVELRVNRRKRPAREDRGLTGPLQSLVP\$ADG\$DADNCQ\$QIMGW\$FTW	688
QY	654	AR\$DPPLNGTFT\$PI\$GALVAVVGVQ\$CGG\$SLL\$SALLAM\$DVBEG-----	700
Db	689	TPDGIPTLSNITIRIPQOLTMVIGQVCGG\$SLL\$LA\$GEMQ\$K\$GAV\$FW\$SSLP\$DSIG	748
QY	701	-----HVAIK\$VAVYVPOQAMIONDSLRENTL\$GCOL\$E\$EYR\$SVIO\$C\$ALL	747
Db	749	ED\$P\$R\$E\$T\$AD\$D\$D\$IRK\$R\$P\$V\$A\$V\$Q\$K\$P\$W\$L\$N\$A\$V\$E\$N\$II\$E\$S\$P\$N\$K\$R\$Y\$M\$V\$E\$A\$C\$LO	808
QY	748	PDL\$IL\$P\$G\$D\$R\$E\$IG\$E\$K\$V\$N\$IS\$G\$Q\$K\$O\$R\$V\$S\$LA\$R\$V\$S\$N\$AD\$IT\$Y\$LD\$D\$P\$LA\$V\$D\$A\$H\$V\$K\$H\$IF	807
Db	809	PDIDILPHGQOIQIGEBGIN\$S\$G\$Q\$R\$ORIS\$V\$A\$R\$AL\$Y\$O\$H\$AN\$V\$FL\$D\$D\$P\$S\$AL\$D\$H\$LS\$D\$H\$M	868
QY	808	ENVI\$G\$P\$G\$M\$K\$-----N\$K\$T\$R\$L\$Y\$H\$S\$M\$Y\$P\$Q\$V\$V\$IT\$Y\$M\$S\$G\$K\$T\$S\$E\$M\$S\$Y\$O\$E\$LL\$A\$D\$O\$AF	862
Db	869	Q-----AGIEHLR\$D\$D\$K\$T\$V\$V\$L\$H\$K\$Q\$Y\$L\$P\$H\$AD\$W\$IT\$AM\$K\$D\$T\$T\$OR\$B\$T\$LO\$FOR\$Q\$E\$OL	923
QY	863	AEFLRTYASTEEOBDAE\$N\$G\$V\$T\$G\$V\$S\$G\$B\$K\$E\$A\$Q\$O\$E\$N\$G\$M\$L\$Y\$T\$D\$S\$A\$G\$O\$L\$O\$R\$O\$L\$S\$S\$S\$S\$Y\$S\$G	922
Db	924	FEH\$K\$T\$LM\$N-RQD\$E\$LE\$K\$E\$T\$V-----E\$R\$K\$A\$T\$E\$P\$Q\$-----L\$R\$A\$M\$S\$B\$D\$G\$--	965
QY	923	DISH\$H\$N\$T\$A\$E\$L\$Q\$A\$E\$K\$E\$T\$M\$K\$LE\$A\$D\$K\$-----Q\$T\$G\$V\$K\$S\$Y\$M\$D\$Y\$M\$A\$L\$G\$I\$F\$S\$FL	977
Db	966	-----LQO\$E\$E\$E\$E\$E\$A\$A\$E\$E\$D\$N\$LS\$M\$H\$Q\$P\$A\$E\$IP\$W\$R\$A\$C\$A\$Y\$LS\$A\$G\$ILL\$LSL	1014
QY	978	SIF\$E\$M\$C\$N\$H\$V\$A\$LS\$N\$Y\$M\$S\$L\$M\$T\$D\$P\$V-----N\$G\$T\$O\$E\$H\$T-----K\$R\$L\$S\$Y\$Y\$A\$G\$	1023
Db	1015	LVF\$Q\$L\$K\$H\$M\$V\$L\$A\$ID\$Y\$W\$L\$A\$K\$T\$D\$S\$A\$L\$T\$LP\$A\$R\$A\$N\$C\$S\$LS\$Q\$C\$T\$LD\$Q\$Y\$V\$A\$V\$V\$Y\$L\$G\$IS\$G	1074
QY	1024	IS\$G\$O\$A\$V\$G\$S\$M\$A\$V\$IS\$G\$IL\$A\$R\$C\$L\$H\$Y\$D\$L\$H\$IL\$R\$S\$P\$M\$S\$P\$E\$R\$P\$S\$G\$L\$V\$N\$R\$F\$K\$E\$LD\$TV	108
Db	1075	I---V\$L\$C\$V\$T\$S\$V\$Y\$E\$W\$G\$L\$K\$A\$K\$H\$S\$L\$N\$R\$II\$T\$A\$P\$R\$E\$F\$E\$T\$T\$P\$D\$S\$IN\$R\$F\$S\$D\$C\$NTI	1131

Dd	966	-----LÖÖBEEEBEASEEDDNLISMLHQRALPMPACAKYLSAGILLISTL	101.4
Qy	978	SIFLEMCNHSALASNYWLSMTDDPIV-----NGTOEH-----KYRLSYGALG	102.2
Dd	1015	LVSFQLKHMVLVAIDPYWLAKMDSALTFLPPAANCSSJSGCTJDQRYAVAWFTVCISIG	107.4
Qy	1024	ISOGINAFGYSMVVISGILASRCLHYDLHLISLRSPMSFEFRPSGLUNRSKEELDTV	108.8
Dd	1075	I---VLCLVTSVTVEWTGKLKVKAKLLHSLNRITLIPAREFFETTPICSLINRFSSDCNTI	113.1

QY 1084 DSMPEVTKMFMSGLFENVIGACIVILLATPIAIIIPPLGIYFPVQRFYVASSRQKRL 1143
DB 1132 DQHPSTLECLSRSTLLCVSALAAVSYTTPVPLVALPLAIVCYFIQCYFVASRDLQOL 1191
QY 1144 ESVRSRPFVSHFNLTIGVSYTRAFEBEORFIHOSDKVDENQAAVPSYIVANMLAVRL 1203
DB 1192 DDTQLPLSHFAETVEGLTITIRAFREARFOOKLEBTDSSNNTASLFLITANMSEVM 1251
QY 1204 -----ECVNCJVLPAALFAVIS--RHSLASGLVLSVSYSLQVTTYLNMVLRMSSE 1254
DB 1252 ATPLPDEYIGACVCLAAVATVINSLSHRELASAGVGLGYALMVSVYLMWVRNLADM 1311
QY 1255 ETNIVAVERLKEYETEKE-----APWQIOETRAPSSWPOGVREPNYCLARYREDLDF 1308
DB 1312 ELQGAIVRIHGLKTESEYEGGLAPSLI-----PKMPPQGIQIQLNLSVRYDSLSKP 1366
QY 1309 VLRIHNTINGEKGVIVGRGAGKSSLTGLFRINESAGEEIIIDGINIKIGLHIDRF 1368
DB 1367 VLKVNMLISPGQKIGICGRGSGKSPSLAFPRVDFEGHIIIDGIDAKLPLHTLRS 1426
QY 1369 KITTIIPQDPVLLFSSGLMNLDPFQOYSDDEVWTSLELAHLDKDFVSALPDKLDHCAEGE 1428
DB 1427 RLSTIILDPPVLFSTIRFNLDPERKCSDTIMEALEIAQLGVYKALPGJDAITTEGGE 1486
QY 1429 NLSVGRQVCLAPALRKTKILVDEATAVLDLETDDLIOSTRTOPEDCTVLTIAHRL 1488
DB 1487 NFSOGQOLFCLARAFVAKTSIFIMDEATASIDMATENNIQLKVVMTAFADRTVVATIAHV 1546
QY 1489 NTIMDYTRVYTLADKEIOEVGAPSDILOQR--GLFPMAMK 1526
DB 1547 HTILSADLVILVKRGAILEFDPKPEKLSRKDSVFASFVR 1585

RESULT 4

US-11-090-439-26
Sequence 26, Application US/11090439
Publication No. US20050266442A1
GENERAL INFORMATION:
APPLICANT: Squillace, Rachel
APPLICANT: Weiner, Michael P.
TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
FILE REFERENCE: 24318-502
CURRENT APPLICATION NUMBER: US/11/090,439
CURRENT FILING DATE: 2005-03-25
PRIOR APPLICATION NUMBER: 60/556,344
PRIOR FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 62
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 26
LENGTH: 1581
TYPE: PRN
ORGANISM: Homo sapiens
US-11-090-439-26

Query Match 21.0%; Score 2100; DB 7; Length 1581;
Best Local Similarity 31.7%; Pred. No. 1.7e-14;
Matches 524; Conservative 306; Mismatches 567; Indels 254; Gaps 37;

QY 32 CFQNTVLVWVCFYLMACFPFFYL-YLSRHGRGYQMTPLNKTALGCF---LMIYCW 86
DB 26 CFVVALNVVPHVFLFITPILFTIGMGSSQSKVHIHS-----TWLHFPCHNLRWILTF 79
QY 87 ADLFYSFMRERGIFLAVP-----FLVSP-----LIGITTLNATFLQLERKGVSSG 136
DB 80 MLPLVVLCEIABGLISDGVESHHLALYMPAGMAFAAVTSV--YNNITSNFPKLI 137
QY 137 IMLEFWALVCAALASKIMTALKEDAOVLDFFDITF-----YVFSLLILQVLVS 189
DB 138 ALVYVMTLAFI-----TKTIKPVKLDHAIIGSQALFCULGLVILXGMLLVE--VN 188
QY 190 CFSRSPULFSET---IHDPNCPRESSAFL-----SRITFWITGLIVGRQPLEGS 239

DB 189 VIRVRRIYFFKTPREKVPREDLODLGVRFLOPFVNLSKGTYYMMANFITAANKPI--- 245
QY 240 DLWSLNKEDTSEQVVPVLVK---NMKECAKTRKQPVKVYSSKDPAPRESSKYDANEE 296
DB 246 DLRAIGK-----LPIVMAALNTNYRLC----- 269
QY 297 VEALIVKSPQ--KEMNBSLFKLYKTGPFFLMSEFFPAIHDLMMFSGPQILKLIKRY-- 353
DB 270 FDAQVRKDIOGTQARAIWQALSHAFGRRLVLSSTFFRIIDLGLGAPPLCIFGVHDIGK 329
QY 354 -NDTKADMQ-----GYFTVLLFVTACTIOTLVYHQGFHICFVSGMBIKT 397
DB 330 ENDFQPKTOFLGVYFVYSSQEFLANAYVLALELALLORTFLQASVYAIETGIDRG 389
QY 398 AVIGAVYRKA--LVITNSARKSSTVGEIVNLMSVDAORFMDLATYIMNIMSAPLOVIAL 455
DB 390 AIOQKIKNKLMLHSTSLMSNGEMTAGOICNLVADITQNMWFFPLCENLAMPQOIIIGV 449
QY 456 YLLMLNIGPSVLACVAVMTLVPNVAMAMKTKTYQVAHMKSKONRIKLAMEIINGIKVL 515
DB 450 ILLYYIIGVSALIGAAVYIILAPVQYFVATKLSQAQRSTLEYSNERLKQTEMLRGIKL 509
QY 516 KLYAMELAFDKVLAIRQBELKVLKKSAYLSAVGTFTWVCPLVALCTPAVYVTTIDENN 575
DB 510 KLYAMENIFKTRVETTRKKTSLRAPAIYTSIFNNYTAIPLAVALITFVGHVSPFEKA 569
QY 576 ILDAQTAFVSLAIFNILRFPNLIIPWISSIVQASVSLKRLRFLSHELEPDSI----- 630
DB 570 DFPSSVAFASISLPHILVTLPLFLSSVVRSTVAVLVQGLSEFLSABEIREDCAPHEP 629
QY 631 -----ERRPVK-----DGGGTN-SITVRNATPTWA 654
DB 630 TPQGPASKYQAVPLRVNRRKPAEEDCRGLTGPLQSLVPSADGADCCVQIMGYEFTWT 689
QY 655 RSDPPTNGITPFSIPAGALVAVVGVCGKSSLSLSALLAEMDKYEG----- 700
DB 690 PDGIPLSNITIRIPRQQLTWIVGVCGKSSLSLLALGEMQKVSQAVMSSLPDSBIGE 749
QY 701 -----HVAIKGSVAVVPOQAMIONDSIRENILEFGQLEBPYRSVIOACALLP 748
DB 750 DSPERETADTDLIRKKGPAVAVASQKPMNLNATYENNIIPSPNNKRVNIVACSLQIP 809
QY 749 DLEILPSGDRTEIGEEKVNISGGQKQKRVSLARAVYSNADYILFDDPLSAVDAAVKGHIFE 808
DB 810 DIDILPHGDOQTQIGERGINISGGQRORISVARALYQHANVFLDDPSALDIHLSDHLMQ 869
QY 809 NVIGPKGMK-----NTRILIVTHSMGYLPQVUVIYIMSGKISEMSYQELLARDGAF 863
DB 870 -----AGILELRDCKRTVVALVTHKQLLPPADWIIMKDGTIQREGTLKDFORSECOLP 924
QY 864 EPLRTVASTEOQDAENGVTVGVSGPGKEAKQKEMNGMLVTDASAGKOLQROLSSSSSYSGD 923
DB 925 EHKMTILAN-RQDQGLKEITYI-----ERKATEPQG-----LSRAMSSDGL--- 965
QY 924 ISRHNSTAELQKAEKKEETWKLMEADKA-----QTGQVKLSYVMDYMAKIGLFIISLS 978
DB 966 -----LQDEEBEEREAASEEDNLSMHLQAEIIPWRACAYLSAGGILLISL 1015
QY 979 IFLPMCHVSAALNMYLSMTDDPIV-----NTOEHT-----KRLSYGALGI 1024
DB 1016 VFSQLKRMVLAIDYMLAKRTDSALTTPAARNCSLSQECTLDQTYAVAMFYVLSIGI 1075
QY 1025 SQGLAVFGVSAVSIIGGILASRCHVDLHLSILRSPMSFPERTPSGMLVNVPFSKELDPTVD 1084
DB 1076 ---VLCVTSVTVWGTGKAKAKRLHRSILNRILIAIPARFETTPGLGSLNRFSSDCNTID 1132
QY 1085 SMPEVIKMFMSGLFENVIGACIVILLATPIAIIIPPLGIYFPVQRFYVASSRQKRL 1144
DB 1133 QHISTLECLSRSTLLCVSALAAVSYTTPVPLVALPLAIVCYFIQCYFVASRDLQOL 1192
QY 1145 SVRSRPFVSHFNLTIGVSYTRAFEBEORFIHOSDKVDENQAAVPSYIVANMLAVRL 1204
DB 1193 DDTQLPLSHFAETVEGLTITIRAFREARFOOKLEBTDSSNNTASLFLITANMSEVM 1252


```
Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 969
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-969

Query Match      9.7%; Score 969; DB 6; Length 410;
Best Local Similarity 50.8%; Pred. No. 1.4e-63;
Matches 197; Conservative 66; Mismatches 121; Indels 4; Gaps 2;

QY 495 MKSKDNRIKLNELINGIKVILKYLAMELAFKDKVLAIRQEBELKYLKKSAYLSAVGTPTWV 554
DB 1 MRQDSARLRTSILRNSKTIKFGHMGCAFIDRVGLRGQELGALRTSGLLFSVLSVFG 60
QY 555 CTPELVALCTFAVAVTIDENNILDAQTAFVSLAFNLFPPLNLPWVSISSIVQASVSLK 614
DB 61 VSTFLVALVFAVHTLVAE-NANNAEKAFVTLVTNLINKAQAFLEPFIHSLQVARSFD 119
QY 615 RLRLFLSHEELPEPSIERPRPKDGGTNSITVRNATFWARSDPTTLNGITFSIPEGALV 674
DB 120 RLVLFLCLEVDPGVVDSSSGSAAGKDCITIQSTFAMSQESPCLIRIMLTPQGLL 179
QY 675 AVGVGVGGKSSLSALIAEMDKYEGHVAIKGSVAVYPOQAMIONDSRENILLEGCOLE 734
DB 180 AVGVGVGGKSSLSALIAEMDKYEGHVAIKGSVAVYPOQAMIONDSRENILLEGCOLE 734
QY 735 PYRYSVIOACALLPDLLEIPSGDRTEIGEKVNLSSGQKQSVSLARAVYSNADITLFPDP 794
DB 240 PMLERVLACALQDPVDSFPEGHTSIIEGQGMNLSGGKQSLSLARAVYRAAYLLDDP 299
QY 795 LSAVDAYGKIFENYIGPKMLKNTKRIILVTHSNSTYPOVDVITVWSGKISEKSTQ 854
DB 300 LAAIDAYGQHVFQVIGPGILQGTTRILVTHLILPQADWITVLNGAIAEMGSYOE 359
QY 855 LLAADGAFAEFLRTYASTEOEDAEENG 882
DB 360 LLQKKGALVCLL---DQARQPGDRGE 384

RESULT 7
US-10-613-744-14
; Sequence 14, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: Mackinnon, Roderick
; APPLICANT: The Rockefeller University
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 14
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-613-744-14

Query Match      8.8%; Score 881.5; DB 6; Length 407;
Best Local Similarity 46.1%; Pred. No. 3.6e-57;
Matches 167; Conservative 75; Mismatches 95; Indels 25; Gaps 4;

QY 1551 LTRLAEPDAPRY----RTERRARFYSKGCNCVAKNTRQGRFLQDVPTTLVLDLKW 1606
DB 16 LPKQARDL--PRHSRDRTRKIQRYVRKQKCNVHGNVRETRYRLDITFTLVLDLKW 74
QY 1607 HTLLIFTFMSFLCSWLLFAMVWMLIAFAHGDLYAYMERGITDLPAPYDVPDAGEGTNP 1666
DB 75 FNLLIFFMVYTVTMLFFGMIWMLIAYIRGDM-----DHIEDSPWT 115
QY 1667 CVTSHSFSSAFLEISIVQVITFGGSMVTEBCPLAILLVQNIQVGLMNAIMLGCIFM 1726
DB 116 CVTNLNGFVASFLPSITETTLIGYGVYITDKCEGIIILLIQVLSIVAFVWGCMFV 175
QY 1727 KTAQAHRAETLFRSKAVITLHGRICFMTLRVGDLEKSMISTITMQRVKTTSRGE 1786
DB 176 KISQPKRAETLVSTAVISMARDKLCIMFRVGDLEKSMISTITMQRVKTTSRGE 235
QY 1787 VVPLHQVDIPMENGVNGGIFLVAFLIIVYIDNSPLDYDAPADLHHODLEIIVLEG 1846
DB 236 FIPLNQTDINVGYYTGDRLEFLVSLIISHINQSPFWELSKQL-PKELELIVILEG 294
QY 1847 VVETTTITTOARTSYLADLEIMGQRFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQL 1906
DB 295 MVEATGMTCQARSYSYIISEILMGYRFPVLTLEDGFYEVVDNSPHEYTESTPISAKEL 354
QY 1907 DE 1908
DB 355 AE 356

RESULT 8
US-10-995-561-967
; Sequence 967, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; APPLICANT: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 967
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-967

Query Match      8.0%; Score 799; DB 6; Length 268;
Best Local Similarity 56.3%; Pred. No. 2.1e-51;
Matches 151; Conservative 56; Mismatches 61; Indels 0; Gaps 0;

QY 1264 LKESSETEKAPMQIOETAPSSWPQYGRVFRFYCCRYREDLDFVLRHINVTINGEKV 1323
DB 1 MODYAMTPKEAPWRLPTCAQPPMPQGGQLEFRDFGLRYRPELPLAVQGVSEFKIHAGSKV 60
QY 1324 GIVRTAGKSSLTIGLFRINESAGEIIGIDGIAIKGLHDLRFKTIIPDPVLFSGS 1383
DB 61 GIVRTAGKSSLTASGLRLQDEALGGIWDGVIAHVGLHTLSRSIITIQDILPFGS 120
QY 1384 LRNWLDPSSQYSDSEVWTSLELAHLKDFVSALPKLHDECAEGENISVGGQRLVCLARA 1443
```

Db 121 LRNNLLDQSHSDRAIWALETVOALKALVSLPQLOKCKADRGEDLSVQKOLLCLARA 180
Qy 1444 LRKTKLVDEATPAVDLETDLIQSTIRQFEDCTVLTAAHRLNTIMDTRIYVLDKG 1503
Db 181 LRRTQLLIDDEATPAVDPGTELOQMAMLSWFAQCTVLTAAHRLSRVMDCAVLDMDKG 240
Qy 1504 EIOEYGAASDLDQORGLFYSMADAGLV 1531
Db 241 QVAESGSPAQLAQKGLFYRLAQESGLV 268

RESULT 9

US-10-613-744-13
; Sequence 13, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: Mackinnon, Roderick
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-613-744-13

Query Match 7.9%; Score 788.5; DB 6; Length 391;
Best Local Similarity 43.7%; Pred. No. 2.2e-50;
Matches 149; Conservative 77; Mismatches 92; Indels 23; Gaps 5;

Qy 1564 RTERRARAFYKSKGNCVNAKINRQGRFL--QDVFITVLDKMPRTLLFTMSFLCSWL 1621
Db 35 RSRQ-RARLVSKGRCNIEFGVNDASQRFIFVDIWTVDLTKRXYKMTVFITAFIGSWF 93
Qy 1622 LFAMVWMLIFAGDLYAVYMEKGITDLAPYRVDVDAEGSTVPCVTSIHSPSSAFLFS 1681
Db 94 LFLGLVYVAVYAKDL-----PEFYPPDNRTPCVENINMGMTSAFLFS 135
Qy 1682 IEVOVITIGEGRNVTCEPLAILILIVONIGMINAIMGCIEMKTAQAHRAETLLIFS 1741
Db 136 LETOVITIGVFRFTEGCATAIFLLIFQSILGVINSFMCAGILLAKISRKXAKITIFS 195
Qy 1742 KHAVITLIRGRCLFPMARVGLDKSMIISATIHQVVRKTSPEGEVPLVHQVDIPMEGV 1801
Db 196 KNAVIVKRGKGLCLLRVALRKSLIGSHIYKGLKTTITPEGETIILDQTNINEFVDA 255
Qy 1802 GNGSIFLVAFLIYHYIDNSPLYDLAPSDLHHODEIIVLEGVETGITTOARTSY 1861
Db 256 GNEMLFISPLTITVYHIIIDHNSPFHMAAETL--SQODELVVFLDGVVESTSATCQVRTSY 314
Qy 1862 LADIELMGQFVPIVAE-EDGRYSVDYSKFGNTIKVPTPLC 1901
Db 315 VPBEVLWGYRFFVIVSKTKGKRVDFHNGKTVETVETPHC 355

RESULT 10

US-10-793-626-326
; Sequence 326, Application US/10793626
; Publication No. US20050255478A1

; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 326
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-326

Query Match 5.1%; Score 514.5; DB 6; Length 548;
Best Local Similarity 27.1%; Pred. No. 4.3e-30;
Matches 157; Conservative 110; Mismatches 211; Indels 101; Gaps 16;

Qy 1002 DPIVNGTOEHTKRLSVYGAIGISQGIAPVGYMAVSGIGILASRCLHVDLSIRSPM 1061
Db 12 DGVIN--NHSLTNOQEFSLHVAIGIALF-----IFLIYVPI 47
Qy 1062 SFEF---RTPSGNLVNRFSKEL-----DTVDSMIPEYIK-----MFWG 1096
Db 48 EFRQYLAQWTSKILYIDIRKQLYNHLQALSVFVANNQGVYISRVINVEQTKDILT 107
Qy 1097 SLFNVIGACIVILLAPI-----AAITIPGL--TYFVQRFYVASSRQKLE 1144
Db 108 GLNNIMWDCITIIIASISMEFLDVKLTFAIFIFPYIITVYFEFGRL-----RKLTVR 162
Qy 1145 SVSRSPVYSHFNETLGVSYIRAFEEQERTIHQSDLKVDENQKAYPSIVANMLAVRLE 1204
Db 163 SOLAEVQGFHERRVOGMSYISFAIEDNEAKNF--NNKNKPLORAFQHTRWAVSFA 219
Qy 1205 CV-----GNCIVLFAALFAVTSRHSLSAGLVGSYSIQVVTYINMLVRMSSEMETNI 1258
Db 220 AINTVTDLGRITVYIGVGSYLAIR--GSITVGTLLAFVGLYELQGLPRLRVSSFTTLQSF 278
Qy 1259 VAVERLKEYSETKE-----APWQIOETAPSSWPQVGRVFNRYCLARYEDLDVLR 1311
Db 279 ASMDRVFQLMDEQYDIDKNGIGAPRISK-----GQIDLKVSFRTYENNEKEVLIH 328
Qy 1312 HINVTINGEKVGIVRTGAGKSLTGLGRINESAGEIIDIGIYIAKIGLHDLRPKT 1371
Db 329 DINTINKGETTAVFVMSGGKSTLINLIPFYDVTQGEILIDHNVKDFLTGSLNRQIG 388
Qy 1372 IIPQDVLFSGLRNNL---DPSQYSDEEVTWTSLEHLKDFVSLPDLKDECAEGE 1428
Db 389 LVQQNDILFSDTKENIILGRF--DATDEVEVAAGANAHDPISLPLNGYDVEVERGV 446
Qy 1429 NUSVQROVLCLARALLRRTKILVDEATPAVDLETDLIQSTIRQFEDCTVLTAAHRL 1488
Db 447 KISGQOKRSLIARILFNNPVLILDEATSAIDLESEAIIOEALDVLISDRQRTLLIAHRL 506
Qy 1489 NITIMDTRIYVLDKGEIOEYGAASDLDQORGLF---YSM 1524
Db 507 STTHADRIVVMEGRIVETGTHQOLINRGAVEHLYSI 545

RESULT 11

US-10-467-657-1346
; Sequence 1346, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega

Best Local Similarity 25.2%; Pred. No. 7e-23;
Matches 149; Conservative 130; Mismatches 234; Indels 79; Gaps 21;

```
QY 977 LSIFLFW--CNHVGALASNYLSTLWDDPIVNGTOEHKTVLSYVGALGISQGIIVFGYS 1034
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 247 LSLFLVLVLSLSEEMALPFETGRITDWIIDQSGADTFRNLITLMSITLTASAVLER--- 303
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 1035 MAVSIG-GILASRCLHY-----DLHSILRSFMSFFERTPSGNLVNRFSEKELPTVDSMI 1087
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 304 ---VGDIYNNVTNGHVSHLQGEVFGAVLRQETEFQONQNGINMSRVETEDSTLSDSL 359
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 1088 PEVTKMFMGSLFNVIYGIIV-IL-ATPIAIIIPPLGLIYFFVQRPYVASSRQK 1141
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 360 SENLSTFLMWYL--VRGCLLGIMLMGVSLSLTMVTLITLPLFLPKVKKMYQLLEVOVR 417
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 1142 RLSEVSRSPVYSHNETLIGSVIRAFEOERFIHQSDLKVDE---NQK---AYYPSIV 1194
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 418 --ESLAKSSQVA--IEALSAMPYTRSPANEGEAKRREKLOETIKLNQKENVAY---A 469
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 1195 ANRW-----IAVRLECVGNCIYLPALPAVISRHSLSAGLVGLSVYSLSQVTTYLNM 1246
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 470 VNSWTTISIGMLTKVGLIYIGQLVTSGA-----VSSGNLVTFLVLYQMQFTQAVEV 520
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 1247 LVKMSSEMETIVAVERLKEVSETEKEAPMOIQETAPSS---WPQVGRVFPKNYCLRY 1302
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 521 LLSIYPRVQKAVGSSEKIFEYLDRTPR-----CPPSGLLTPPLHLEGLVQFQDVSPAY 572
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 1303 --REDLPVLRHINVTINGEKGIVGRGTGAKSLSLTGLFRINSAEGEIIIDGINIAK 1360
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 573 PNRPDV-LVLOGLFTTLRFGVETALVGPNGSKSTVALNLONLYOPTGGQLLDGKPLPQ 631
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 1361 IGLHDLRFKTIIPQDPVLPFSGSLRML--DPFSQYSDDEVWTSLELAHLKDFVSALPDK 1418
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 632 YEHRYLHRQVAAVQEPQVFGRLQENIAYGLTQKPTWEIITAAAVKSGAHSFISGLPQG 691
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 1419 LDHECARGENTLSVGQQLVCLARALLKTKILVIDENTAVID---LETDDLQSTIRT 1474
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 692 YDTEVDGSQLSGQQAVALARALIRKPCVLILDDATSAIDANSQLOVEQLLYES--P 749
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 1475 QFEDCTVLTIAHRLNTIMDYTRVYLDKGEIOEGAPSDLQORGLFYSMK 1526
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 750 ERYRSVLLITQHLSLVEQADHILFLBGGALRBGGTHQOLMEKKGCYAMAVQ 801
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
```

Search completed: December 15, 2005, 15:50:18
Job time : 12.3405 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:27:22 / Search time 178.197 Seconds
(without alignments)
4800.712 Million cell updates/sec

Title: US-10-665-283-8

Perfect score: 10016
Sequence: 1 MALRGFCNADGSDPLMDMNV.....RSVAVAKKPKFSISPSLS 1947

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10016	100.0	1947	ADY86945	Ady86945 Human MRP
2	9873	98.6	1927	ADY86938	Ady86938 Human MRP
3	9861	98.5	1927	ADY86943	Ady86943 MRP1-mous
4	9704	96.9	1891	ADY86941	Ady86941 Human MRP
5	7860	78.5	1531	AAW57486	AAW57486 Human MRP
6	7860	78.5	1531	AAW74471	AAW74471 Human mul
7	7860	78.5	1531	AAW98984	AAW98984 Human mul
8	7860	78.5	1531	AAW55799	AAW55799 Human mul
9	7860	78.5	1531	AAW78873	AAW78873 Human mul
10	7860	78.5	1531	ABM35012	ABM35012 Prostate
11	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
12	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
13	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
14	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
15	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
16	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
17	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
18	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
19	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
20	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
21	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
22	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
23	7860	78.5	1531	ABM35012	ABM35012 Cancer ba
24	7860	78.5	1531	ABM35012	ABM35012 Cancer ba

25	7849	78.4	1531	2	AAW98983	AAW98983 Human mul
26	7849	78.4	1531	3	AAW55798	AAW55798 Human mul
27	7849	78.4	1531	3	AAW78872	AAW78872 Human mul
28	7849	78.4	1531	3	AAW03582	AAW03582 Multidrug
29	7763	77.5	1515	8	ADU24094	ADU24094 Human cys
30	7610	76.0	1489	2	AAW96952	AAW96952 Multi-dru
31	7587.5	75.8	1482	8	ADN03902	ADN03902 Antipsoi
32	7587.5	75.8	1482	8	ADP23422	ADP23422 PRO poly
33	7572	75.6	1481	5	ABP52108	ABP52108 Homo sapi
34	7560.5	75.5	1482	5	AAW96953	AAW96953 Multi-dru
35	7535	75.2	1475	9	ADW06206	ADW06206 Cyclin-de
36	7510.5	75.0	1472	9	ADW06204	ADW06204 Cyclin-de
37	7500.5	74.9	1472	2	AAW96954	AAW96954 Multi-dru
38	7230.5	72.2	1530	7	ADP56611	ADP56611 Bovine MR
39	7195.5	71.8	1416	9	ADW06208	ADW06208 Cyclin-de
40	7194	71.8	1417	2	AAW96955	AAW96955 Multi-dru
41	7037	70.3	1388	7	ADW56912	ADW56912 Human hea
42	7002.5	69.9	1528	2	AAW57487	AAW57487 Murine mu
43	7002.5	69.9	1528	2	AAW74472	AAW74472 Mouse mul
44	7002.5	69.9	1528	2	AAW98985	AAW98985 Mouse mul
45	7002.5	69.9	1528	3	AAW55800	AAW55800 Murine mu

ALIGNMENTS

RESULT 1
ID ADY86945 standard; protein; 1947 AA.
XX
AC ADY86945;
XX
DT 02-JUN-2005 (first entry)
XX
DE Human MRP1-mouse Klr6.2-HA fusion protein, SEQ ID NO: 8.
XX
KW Ionophore; biosensor; drug screening; diagnostic;
KW microorganism detection; potassium channel; fusion protein;
KW multidrug resistance protein 1; MRP1; Klr6.2.
XX
OS Homo sapiens.
OS Mus musculus.
OS Chimeric.
OS Unidentified.
XX
XX
XX Key Location/Qualifiers
XX FT 1..1531 /note="Multidrug resistance protein 1 (MRP1)"
XX FT 1532..1537 /note="Hexaglycine spacer"
XX FT Region /note="Klr6.2 protein containing HA epitope peptide"
XX FT 1538..1947 /note="Klr6.2 protein containing HA epitope peptide"
XX FT 1539..1947 /note="Klr6.2 protein containing HA epitope peptide"
XX FT Region /note="HA epitope peptide"
XX
XX US2005063989-A1.
XX
XX 24-MAR-2005.
XX
XX 22-SEP-2003; 2003US-00665283.
XX
XX 22-SEP-2003; 2003US-00665283.
XX
XX (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX
XX Derand R, Garcia B, Prost A, Revilloud J, Vivaudou M;
XX WPI, 2005-252611/26.
XX
XX New ion channel hybrid protein, used as electrical sensor for screening
XX an agonist/antagonist of a membrane protein and for detecting a
XX contaminant/pollutant in a sample.
XX

PS Claim 25; SEQ ID NO 8; 78pp; English.

XX The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 protein containing HA epitope peptide.

XX Sequence 1947 AA;

Query Match 100.0%; Score 10016; DB 9; Length 1947;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTYCFONTVLVWVPCFYLAACPPFFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTYCFONTVLVWVPCFYLAACPPFFLYLSRH 60
QY 61 DRGIQMTPLNKTALGELIMIVCMADLFYSFWRSGIFLAPFVLSPITLGLITLLA 120
DB 61 DRGIQMTPLNKTALGELIMIVCMADLFYSFWRSGIFLAPFVLSPITLGLITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMLVALCALALRSKMTALKEDAQVDLFFDITFYVFS 180
DB 121 TFLIQLERRKGVSSGIMLTFMLVALCALALRSKMTALKEDAQVDLFFDITFYVFS 180
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCESSASFLSRITTWITGLIVRGYROPLEGGSD 240
DB 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCESSASFLSRITTWITGLIVRGYROPLEGGSD 240
QY 241 LMSINKEDTSROVPLVYKMKKKECAKTRKQPVKYVYSSKDPADPKESSKVDANEVVAL 300
DB 241 LMSINKEDTSROVPLVYKMKKKECAKTRKQPVKYVYSSKDPADPKESSKVDANEVVAL 300
QY 301 IVKSPQKEMNPSLFEKVLKYTFGPFYFLMSFFPKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFEKVLKYTFGPFYFLMSFFPKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
QY 361 WQGFYFVLLFVYACLOTVLHOFPHICFVSGMRITKIVAGVYRKALVITNSARKSTV 420
DB 361 WQGFYFVLLFVYACLOTVLHOFPHICFVSGMRITKIVAGVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAORMDLATYINMTWSAPLOYITLALYLLMLNGSPVLAVGAVMVLMPVN 480
DB 421 GEIVNLSVDAORMDLATYINMTWSAPLOYITLALYLLMLNGSPVLAVGAVMVLMPVN 480
QY 481 AVMAKTKTYOVAAHMSKDNRIKLMNEILNGIKVULKYAMELAFDKVLAIROBELKVLK 540
DB 481 AVMAKTKTYOVAAHMSKDNRIKLMNEILNGIKVULKYAMELAFDKVLAIROBELKVLK 540
QY 541 KSAVLAVGTFETWCTPEFLVACTFAVYVITDENNIIDAOQAFLVSLFNLIRPPLNLP 600
DB 541 KSAVLAVGTFETWCTPEFLVACTFAVYVITDENNIIDAOQAFLVSLFNLIRPPLNLP 600
QY 601 MVISISYQASVSLKRLRIFLSHELEPDSIERRPYKDGCTNSITVRNATTTMARSDPT 660
DB 601 MVISISYQASVSLKRLRIFLSHELEPDSIERRPYKDGCTNSITVRNATTTMARSDPT 660
QY 661 LINGTFSPREGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAKGSAAVYPOQAMIOND 720
DB 661 LINGTFSPREGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAKGSAAVYPOQAMIOND 720
QY 721 SLRENILFEGCOLLEBYRSVIAQCALLPDLBILPSGDRTEIGEKGVLNLSGQKORVSLAR 780
DB 721 SLRENILFEGCOLLEBYRSVIAQCALLPDLBILPSGDRTEIGEKGVLNLSGQKORVSLAR 780
QY 781 AVYNSADIIYLFDDPLSANDAVHGKAI FENVIGPKMKLKNKTRILVTHSMSTLPQVDVIV 840
DB 781 AVYNSADIIYLFDDPLSANDAVHGKAI FENVIGPKMKLKNKTRILVTHSMSTLPQVDVIV 840

QY 841 MSGGKISEMGSYOELLARDGAFAELFTYASTROEDQDAEENGVTGVSQPKGKAXOMENGM 900
DB 841 MSGGKISEMGSYOELLARDGAFAELFTYASTROEDQDAEENGVTGVSQPKGKAXOMENGM 900
QY 901 LVYDSACKQIQORQLSSSSSSSISGDISRHHNSTAELQKAEKKEFTWKLMEADKATQGVKL 960
DB 901 LVYDSACKQIQORQLSSSSSSSISGDISRHHNSTAELQKAEKKEFTWKLMEADKATQGVKL 960
QY 961 SVYDWYKKAIGLFSFISITLFMCNHSALASNWMLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYDWYKKAIGLFSFISITLFMCNHSALASNWMLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQGIAVFGYSMAVSIIGIILASRCLAVDLHLSILRSPMSFEPTPSGNLVNRSKEL 1080
DB 1021 ALGISQGIAVFGYSMAVSIIGIILASRCLAVDLHLSILRSPMSFEPTPSGNLVNRSKEL 1080
QY 1081 DTYDSMTPEVITKPMGSLFENVIGACIYIILATPIAIIIPPLGIYFVQRFYVASSROL 1140
DB 1081 DTYDSMTPEVITKPMGSLFENVIGACIYIILATPIAIIIPPLGIYFVQRFYVASSROL 1140
QY 1141 KRLESVRSPPYSHFNETLLGVSIVIRAFEOERFIHOSDLKVDENOKAYPSIVANBMLA 1200
DB 1141 KRLESVRSPPYSHFNETLLGVSIVIRAFEOERFIHOSDLKVDENOKAYPSIVANBMLA 1200
QY 1201 VRLCEVNCIYVFAALPAVVISRHSLSAGLVGLSVSYSLQYTYILNMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYVFAALPAVVISRHSLSAGLVGLSVSYSLQYTYILNMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSERTEKAPMIOETAPPSWPOYGRVFRNYCLARYDDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSERTEKAPMIOETAPPSWPOYGRVFRNYCLARYDDLPVLRHINVTINGG 1320
QY 1321 EKVGI VERTGAGKSLTGLFRINESAGEIIDGINIAIKGLHDLKFKITIIPODPVLF 1380
DB 1321 EKVGI VERTGAGKSLTGLFRINESAGEIIDGINIAIKGLHDLKFKITIIPODPVLF 1380
QY 1381 SGLSRNMLDPFSQYSDSEVWTSLELAHKDFVSAI.PDKLDHECAGEGEMLSVGQROLVCL 1440
DB 1381 SGLSRNMLDPFSQYSDSEVWTSLELAHKDFVSAI.PDKLDHECAGEGEMLSVGQROLVCL 1440
QY 1441 ARALIRKTKILVDEATAVNDLETDLIIOSTIRQOFEDCVYLTIAHLNLTIMYTRIVYL 1500
DB 1441 ARALIRKTKILVDEATAVNDLETDLIIOSTIRQOFEDCVYLTIAHLNLTIMYTRIVYL 1500
QY 1501 DKGEIOEYGA.PSDILOORG.FYSMAKDAGLVGGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
DB 1501 DKGEIOEYGA.PSDILOORG.FYSMAKDAGLVGGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
QY 1561 PRYTRRRRARFVSKKNCNVAAHKNIREOGFLODVFETTLVLDLKM.PHTLLIFTMSFLCSW 1620
DB 1561 PRYTRRRRARFVSKKNCNVAAHKNIREOGFLODVFETTLVLDLKM.PHTLLIFTMSFLCSW 1620
QY 1621 LLEPAMVWMLIAFAHGDLYAYMEKGITDLAPYDVDPYABEGINVCVCTSHSSSAFLF 1680
DB 1621 LLEPAMVWMLIAFAHGDLYAYMEKGITDLAPYDVDPYABEGINVCVCTSHSSSAFLF 1680
QY 1681 SIEVOVITIGCGRMVTECEPLAIIILIVONIVGMINAIVMGCIFEMTAQAHRRAETLIF 1740
DB 1681 SIEVOVITIGCGRMVTECEPLAIIILIVONIVGMINAIVMGCIFEMTAQAHRRAETLIF 1740
QY 1741 SKHAVITLRHGRLCFMLRVGDLRKSMIISATIHQVVRKTTSPSGEVVPLPLHQVDIPENG 1800
DB 1741 SKHAVITLRHGRLCFMLRVGDLRKSMIISATIHQVVRKTTSPSGEVVPLPLHQVDIPENG 1800
QY 1801 VGGNGIFLVAELIYVHVIDSNSPLYLAPSDLHHHODEIIVILEGVETTGITTOARTS 1860
DB 1801 VGGNGIFLVAELIYVHVIDSNSPLYLAPSDLHHHODEIIVILEGVETTGITTOARTS 1860
QY 1861 YLADEIILMGORFVYIVAEEOGRYSDVYSKRENTIKVPLPCTARQLEDRLDALTLAS 1920
DB 1861 YLADEIILMGORFVYIVAEEOGRYSDVYSKRENTIKVPLPCTARQLEDRLDALTLAS 1920

QY 1921 SRGFLRRKRSVAVAKAKPKFSISPDLS 1947
 DB 1921 SRGFLRRKRSVAVAKAKPKFSISPDLS 1947

RESULT 2
 ID ADY86938 standard; protein; 1927 AA.
 AC ADY86938;

XX 02-JUN-2005 (first entry)
 DE Human MRP1-mouse Klr6.2 fusion protein, SEQ ID NO: 1.

XX Ionophore; biosensor; drug screening; diagnostic;
 KM microorganism detection; potassium channel; fusion protein;
 KW multidrug resistance protein 1; MRP1; Klr6.2.

XX Homo sapiens.
 OS Mus musculus.
 OS Chimeric.

XX Key Location/Qualifiers
 FH 1.1531
 FT /note= "Multidrug resistance protein 1 (MRP1)"
 FT 1532..1537
 FT /note= "Hexaglycine spacer"
 FT 1538..1927
 FT Region /note= "Klr6.2 protein"

XX US2005063989-A1.

XX 24-MAR-2005.

XX 22-SEP-2003; 2003US-00665283.

XX 22-SEP-2003; 2003US-00665283.

XX (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

XX Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
 DR WPI; 2005-252611/26.

XX New ion channel hybrid protein, used as electrical sensor for screening
 PT an agonist/antagonist of a membrane protein and for detecting a
 PT contaminant/pollutant in a sample.

XX Claim 25; SEQ ID NO 1; 78pp; English.

XX The present invention relates to a hybrid protein consisting essentially
 CC of the fusion of a membrane protein with an ion channel which is not
 CC naturally coupled to the membrane protein. The hybrid protein is used as
 CC an electrical sensor for screening of an agonist/antagonist of a membrane
 CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
 CC invention is also useful for medical diagnostics and microorganism
 CC detection. The present sequence is a fusion protein comprising human
 CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
 CC channel Klr6.2 protein.

XX Sequence 1927 AA;

Query Match 98.6%; Score 9873; DB 9; Length 1927;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1927; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

QY 1 MALRGFCSADGSPILMDMNTANTSNPDFKCFONTLVVPCFYLMACPFYFLYSRH 60
 DB 1 MALRGFCSADGSPILMDMNTANTSNPDFKCFONTLVVPCFYLMACPFYFLYSRH 60
 QY 61 DRGYIQMTPLNKTKTALGFLLMIIVCWADLFYFWEERSRGIFLAPVFLVSPFTLLGITTLA 120

DB 61 DRGYIQMTPLNKTKTALGFLLMIIVCWADLFYFWEERSRGIFLAPVFLVSPFTLLGITTLA 120
 QY 121 TPLIQLEERRRGVSSGIMLTFFMVALVCALAIIRSKITMTALKEDAVDLRDTIFYYYS 180
 DB 121 TPLIQLEERRRGVSSGIMLTFFMVALVCALAIIRSKITMTALKEDAVDLRDTIFYYYS 180
 QY 181 LLLIQVLVSCFSSRSLFSETHIDPNCPESSASFLSRITFMTGTIIVGYROPLGSD 240
 DB 181 LLLIQVLVSCFSSRSLFSETHIDPNCPESSASFLSRITFMTGTIIVGYROPLGSD 240
 QY 241 LMSLNKEDTSEQVVPVLYKNMKKCAKTRKQPVKVVYSKDPAPCKESSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLYKNMKKCAKTRKQPVKVVYSKDPAPCKESSKVDANEVEAL 300
 QY 301 IVKSPQKEMNPSPFLKYLKTFPGYFLMSFFPKAIHDLMSGQIILKLIKFNVDTPAD 360
 DB 301 IVKSPQKEMNPSPFLKYLKTFPGYFLMSFFPKAIHDLMSGQIILKLIKFNVDTPAD 360
 QY 361 MOGYFTVLLEFVRAQOTLVHOFPHICFVSGMRKTAIVGAYRKALVTNSARKSSTV 420
 DB 361 MOGYFTVLLEFVRAQOTLVHOFPHICFVSGMRKTAIVGAYRKALVTNSARKSSTV 420
 QY 421 GEIVNIMSVDQRFMDLATYINMIWSAPLOVILALYLMNLGSPVLGAVAVMTAMVPV 480
 DB 421 GEIVNIMSVDQRFMDLATYINMIWSAPLOVILALYLMNLGSPVLGAVAVMTAMVPV 480
 QY 481 AVNARKTKTYQVAHMSKONRIKLANEILNGIKYLVKLYAMELAPKOKVLAIRBELKVLK 540
 DB 481 AVNARKTKTYQVAHMSKONRIKLANEILNGIKYLVKLYAMELAPKOKVLAIRBELKVLK 540
 QY 541 KSAVYLSAVGFTFVNCPELVALCTPAVYVITIDNNILDAQTAVSLAFNIIILPPLNLP 600
 DB 541 KSAVYLSAVGFTFVNCPELVALCTPAVYVITIDNNILDAQTAVSLAFNIIILPPLNLP 600
 QY 601 MVISSIVQASVSLKRLRIFLSHELEBPSIERRPVKGSGTNSITYNATFTWARSBPPT 660
 DB 601 MVISSIVQASVSLKRLRIFLSHELEBPSIERRPVKGSGTNSITYNATFTWARSBPPT 660
 QY 661 LMGITTSIPRGALVAVGVGCGKSSILSLALEMKRVESHVAIKGSVAVYPOQAWIOND 720
 DB 661 LMGITTSIPRGALVAVGVGCGKSSILSLALEMKRVESHVAIKGSVAVYPOQAWIOND 720
 QY 721 SLRENIIFGQLEPEPYRSTYQACALLPDLEILPQSDRETEIGKGVNLSSGQORVSLAR 780
 DB 721 SLRENIIFGQLEPEPYRSTYQACALLPDLEILPQSDRETEIGKGVNLSSGQORVSLAR 780
 QY 781 AVYSNADIVLFDDPLSLAVDAHVAGKHIFENVYIGPKMLKNKTRILVTHSMYSYLVQVDVYIV 840
 DB 781 AVYSNADIVLFDDPLSLAVDAHVAGKHIFENVYIGPKMLKNKTRILVTHSMYSYLVQVDVYIV 840
 QY 841 MSGGKISEMGSYQELLARDGAPAEPLRTYASTBOEDAEENGVTGVSQPKKAKOMENGM 900
 DB 841 MSGGKISEMGSYQELLARDGAPAEPLRTYASTBOEDAEENGVTGVSQPKKAKOMENGM 900
 QY 901 LVTDSAGKQORLSSSSSVSGDISRRHNSTALQKAEAKKETWLMENDKQOTQVYL 960
 DB 901 LVTDSAGKQORLSSSSSVSGDISRRHNSTALQKAEAKKETWLMENDKQOTQVYL 960
 QY 961 SVYWDYMKAIQGLIFSLISFLFWCNHVSALASAYMLSLMTDDIYVNGTOEHTYRLSVYG 1020
 DB 961 SVYWDYMKAIQGLIFSLISFLFWCNHVSALASAYMLSLMTDDIYVNGTOEHTYRLSVYG 1020
 QY 1021 ALGISQGIIVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFPERTPSGNLVNRFSKEL 1080
 DB 1021 ALGISQGIIVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFPERTPSGNLVNRFSKEL 1080
 QY 1081 DTVDSMTPEVITKFMKWSLFRNVIGACVILLATPAAIILPGLIYFPQRFVASSROL 1140
 DB 1081 DTVDSMTPEVITKFMKWSLFRNVIGACVILLATPAAIILPGLIYFPQRFVASSROL 1140
 QY 1141 KRLSEVSRSPVYSHFETLLGVSVIRAFEOERFIHQSLDKVDENOKAYVPSIVANRWLA 1200
 DB 1141 KRLSEVSRSPVYSHFETLLGVSVIRAFEOERFIHQSLDKVDENOKAYVPSIVANRWLA 1200

```

QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSSEMETNIVA 1260
QY 1261 VERLKESEPTKEKPMQIQETAPSPSSWPQVGRVERFRNYCLARYREDLPVLAHINVTINGG 1320
Db 1261 VERLKESEPTKEKPMQIQETAPSPSSWPQVGRVERFRNYCLARYREDLPVLAHINVTINGG 1320
QY 1321 EKVGIVERTGAGKSLTGLFRINESAGEIIIGINIAKIGLHDLRKITIIIPQDPLF 1380
Db 1321 EKVGIVERTGAGKSLTGLFRINESAGEIIIGINIAKIGLHDLRKITIIIPQDPLF 1380
QY 1381 SGLIRNMLDPFSPQYSDSEWMTSLBLAHLDKDFVSALPDRLDHECAEGENLSVGQRQLVCL 1440
Db 1381 SGLIRNMLDPFSPQYSDSEWMTSLBLAHLDKDFVSALPDRLDHECAEGENLSVGQRQLVCL 1440
QY 1441 ARALLRKTILVDEBATAVDLETDLIQSTIRQFECTVLTIAHRLNTIMDTTRIVL 1500
Db 1441 ARALLRKTILVDEBATAVDLETDLIQSTIRQFECTVLTIAHRLNTIMDTTRIVL 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYMAKADAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Db 1501 DKGEIOEYGAPSDILQORGLFYMAKADAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
QY 1561 PRYTRERRARFVSKGNCVNAHKNIRBQGRFLQDVFTTLVDLKPHTLLIFMSFLCSW 1620
Db 1561 PRYTRERRARFVSKGNCVNAHKNIRBQGRFLQDVFTTLVDLKPHTLLIFMSFLCSW 1620
QY 1621 LLEFAMWMLTFAFGDLVAYMEKGTDLAPYRVDYDAGSGTNVPCCTSIHSSSALFL 1680
Db 1621 LLEFAMWMLTFAFGDLVAYMEKGTDLAPYRVDYDAGSGTNVPCCTSIHSSSALFL 1680
QY 1681 SIEVOVTIGFGSMYTECPALILILYQNTVGLMINIMGCTFMKTAQARRAETLIIF 1740
Db 1681 SIEVOVTIGFGSMYTECPALILILYQNTVGLMINIMGCTFMKTAQARRAETLIIF 1740
QY 1741 SKHAVITLHRGLCFMRLRVGDLRKSMTISATIHQVVKTTSPBEGVPLHQVDIPMENG 1800
Db 1741 SKHAVITLHRGLCFMRLRVGDLRKSMTISATIHQVVKTTSPBEGVPLHQVDIPMENG 1800
QY 1801 VGGNGIFLVAALIIYHVIDSNSPLYDLAPSDLHHQDLEIIVILEGVETTGITTOATS 1860
Db 1801 VGGNGIFLVAALIIYHVIDSNSPLYDLAPSDLHHQDLEIIVILEGVETTGITTOATS 1860
QY 1861 YLADEIILMGORFPIVABEDGRYSVDYSGFNTIKVPTPLCTAROLDERSILALTLAS 1920
Db 1861 YLADEIILMGORFPIVABEDGRYSVDYSGFNTIKVPTPLCTAROLDERSILALTLAS 1920
QY 1921 SRGPILRKRSVAVAKAKPKFSISPDLS 1947
Db 1921 SRGPILRKRSVAVAKAKPKFSISPDLS 1947
QY 1901 SRGPILRKRSVAVAKAKPKFSISPDLS 1927
Db 1901 SRGPILRKRSVAVAKAKPKFSISPDLS 1927

RESULT 3
ADY86943
ID ADY86943 standard; protein; 1927 AA.
XX
AC ADY86943;
XX
DT 02-JUN-2005 (first entry)
XX
DE MRP1-mouse Klr6.2 mutant (KR370AA) fusion protein, SEQ ID NO: 6.
XX
KW Ionophore; biosensor; drug screening; diagnostic;
KW microorganism detection; potassium channel; fusion protein;
KW multidrug resistance protein 1; MRP1; Klr6.2; muttein.
XX
OS Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX
FH Key Location/Qualifiers

```

```

FT Region
FT 1. .1531
FT /note= "Multidrug resistance protein 1 (MRP1)"
FT Region
FT 1532. .1537
FT /note= "Hexaglycine spacer"
FT Region
FT 1538. .1927
FT /note= "Klr6.2 mutant protein"
FT Misc-difference
FT 1907
FT /note= "Wild-type Iys substituted by Ala"
FT Misc-difference
FT 1908
FT /note= "Wild-type Arg substituted by Ala"
FT US2005063989-A1.
FT 24-MAR-2005.
FT 22-SEP-2003; 2003US-00665283.
FT 22-SEP-2003; 2003US-00665283.
FT 22-SEP-2003; 2003US-00665283.
FT (COMS ) COMMISSARIAT ENERGIE ATOMIQUE.
FT Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
FT WPI; 2005-252611/26.
FT
FT New ion channel hybrid protein, used as electrical sensor for screening
FT an agonist/antagonist of a membrane protein and for detecting a
FT contaminant/pollutant in a sample.
FT
FT Claim 25; SEQ ID NO 6; 78bp; English.
XX
XX The present invention relates to a hybrid protein consisting essentially
XX of the fusion of a membrane protein with an ion channel which is used
XX naturally coupled to the membrane protein. The hybrid protein is used as
XX an electrical sensor for screening of an agonist/antagonist of a membrane
XX protein, drugs and for detecting a contaminant/pollutant in a sample. The
XX invention is also useful for medical diagnostics and microorganism
XX detection. The present sequence is a fusion protein comprising human
XX multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
XX channel Klr6.2 mutant (KR370AA) protein.
XX
XX Sequence 1927 AA.
SQ
Query Match 98.5%; Score 9861; DB 9; Length 1927;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 20; Gaps 2;
QY 1 MALRGFSGADGSDPLMDMNTWTMTNSNDFTKCFQNTVLYVWPCTYLMACFPFFLYLSRH 60
Db 1 MALRGFSGADGSDPLMDMNTWTMTNSNDFTKCFQNTVLYVWPCTYLMACFPFFLYLSRH 60
QY 61 DRGYIOMTPUNKKTALGFLIMIVCMADLPFSFERSRGIFLAVVPLVSPTLGTTTLA 120
Db 61 DRGYIOMTPUNKKTALGFLIMIVCMADLPFSFERSRGIFLAVVPLVSPTLGTTTLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMLVALVCALALIRSKIMTALKEDAQVDLFFRDIFFYYFS 180
Db 121 TFLIQLERRKGVSSGIMLTFMLVALVCALALIRSKIMTALKEDAQVDLFFRDIFFYYFS 180
QY 181 LLLIQLVISCPSDRSPLFSETIHDNPNCPBSSASFSLRITRWITGLIVRGROPLEBSD 240
Db 181 LLLIQLVISCPSDRSPLFSETIHDNPNCPBSSASFSLRITRWITGLIVRGROPLEBSD 240
QY 241 LMSLNKEDTSQVVPVLYKMKKCAKTRQPVVYVSSQDPAOPKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVLYKMKKCAKTRQPVVYVSSQDPAOPKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLKFVLYKTFGPFYLMSPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLKFVLYKTFGPFYLMSPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360
QY 361 WQGYFTYVLLFVTAQLOTLVIAHQYFHI CFVSGMRIKTAIVAGVYRKALVITNSARKSTV 420

```


Db 361 WOGFYTVLLFVTAQTLVLAHQYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
 Qy 421 GETVNLMSVDAORPMDLATYINMWSAPLOVYTLALYLLMLNGSVTAAGVAVMLAMPVN 480
 Db 421 GETVNLMSVDAORPMDLATYINMWSAPLOVYTLALYLLMLNGSVTAAGVAVMLAMPVN 480
 Qy 481 AVAMKTKTYOVAMHMSKDNRIKLMELINGIKVLYLKLAMELAFKDYKLYI ROBELKYLK 540
 Db 481 AVAMKTKTYOVAMHMSKDNRIKLMELINGIKVLYLKLAMELAFKDYKLYI ROBELKYLK 540
 Qy 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVITDENNNILDAQTAFAVSLFNLRLPPLNLP 600
 Db 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVITDENNNILDAQTAFAVSLFNLRLPPLNLP 600
 Qy 601 MVSIVQASVSLKRLIFLSHELEPDSIERPRVKKGGGNTSTYNNATFTAKRSBPPT 660
 Db 601 MVSIVQASVSLKRLIFLSHELEPDSIERPRVKKGGGNTSTYNNATFTAKRSBPPT 660
 Qy 661 LMGITFESIPGALVAVGVGCGKSSLSALLAEMDKVEGHVAIKGSVAVYPOOAMTQND 720
 Db 661 LMGITFESIPGALVAVGVGCGKSSLSALLAEMDKVEGHVAIKGSVAVYPOOAMTQND 720
 Qy 721 SLRENIILFGQLEBPYRSVYIOACALLPDLIELPSGDRTEIGEXVNLSSGQKQVSLAR 780
 Db 721 SLRENIILFGQLEBPYRSVYIOACALLPDLIELPSGDRTEIGEXVNLSSGQKQVSLAR 780
 Qy 781 AVYNSMDIYLPDPLSLVDAHNGKHIFENYIGPKGMKAKTRILVTHSMYTLPOVDYIIV 840
 Db 781 AVYNSMDIYLPDPLSLVDAHNGKHIFENYIGPKGMKAKTRILVTHSMYTLPOVDYIIV 840
 Qy 841 MSGGKISEMGSYOELLARDGAPAFLERTYASTEOBOAEENGVTGVS GPKKAKOMENGM 900
 Db 841 MSGGKISEMGSYOELLARDGAPAFLERTYASTEOBOAEENGVTGVS GPKKAKOMENGM 900
 Qy 901 LVYDSAGKQLOROLSSSSSYSGDISRHHNSTAELOKAEKKEETKLMLEADKAQTQGVKL 960
 Db 901 LVYDSAGKQLOROLSSSSSYSGDISRHHNSTAELOKAEKKEETKLMLEADKAQTQGVKL 960
 Qy 961 SVYWDYMKAIGLTSTFLSTFLMCONHYSLASNYWLSMTDDPIVNGTOHTYKRLSVYG 1020
 Db 961 SVYWDYMKAIGLTSTFLSTFLMCONHYSLASNYWLSMTDDPIVNGTOHTYKRLSVYG 1020
 Qy 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLAHVDLHSLIRSPMSFEERPSCNINLRPSKEL 1080
 Db 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLAHVDLHSLIRSPMSFEERPSCNINLRPSKEL 1080
 Qy 1081 DTVDSMIPEVIKMFMSGLFNVIACIYILATPPIAIIIPPLGIYFFVORFYVASSROL 1140
 Db 1081 DTVDSMIPEVIKMFMSGLFNVIACIYILATPPIAIIIPPLGIYFFVORFYVASSROL 1140
 Qy 1141 KRLSVSRSPYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYYPSIVANRWLA 1200
 Db 1141 KRLSVSRSPYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYYPSIVANRWLA 1200
 Qy 1201 VRLCEGNCIVLPAALPAVISRHSLSAGVLSVSYSLQVTTYTLNMLVRRSSEMETNYA 1260
 Db 1201 VRLCEGNCIVLPAALPAVISRHSLSAGVLSVSYSLQVTTYTLNMLVRRSSEMETNYA 1260
 Qy 1261 VERLKEYSETEKAPMOIOETAPSSWPQGVAREFRNYCLARYEDLDVLRHINVTINGG 1320
 Db 1261 VERLKEYSETEKAPMOIOETAPSSWPQGVAREFRNYCLARYEDLDVLRHINVTINGG 1320
 Qy 1321 EKVGIVGRGAGKSSLTGLFRINESABGEIIGDGINIAKIGLHDLRFKTTIIPDPVLF 1380
 Db 1321 EKVGIVGRGAGKSSLTGLFRINESABGEIIGDGINIAKIGLHDLRFKTTIIPDPVLF 1380
 Qy 1381 SCSLRNMLDPFSQYDEEVTSLLELAHLKDFVSLAPDKLHCEAEGENNSVGOBOLVCL 1440
 Db 1381 SCSLRNMLDPFSQYDEEVTSLLELAHLKDFVSLAPDKLHCEAEGENNSVGOBOLVCL 1440
 Qy 1441 ARALLKRTKILVDEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVYL 1500
 Db 1441 ARALLKRTKILVDEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVYL 1500

Qy 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPREYVLTRLAEPAE 1560
 Db 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPREYVLTRLAEPAE 1560
 Qy 1561 PRYTERRRARFVSKKGNCAVAKHNTIEQGRFLQDVFTTLVDLKM PHTLLI FTWSFLCSW 1620
 Db 1561 PRYTERRRARFVSKKGNCAVAKHNTIEQGRFLQDVFTTLVDLKM PHTLLI FTWSFLCSW 1620
 Qy 1621 LFLPAMVWMLIAFAHGDLYAAMEKGTDLADPYVDVDPVYAGEGTVNPCVTSIHSSSAFLF 1680
 Db 1621 LFLPAMVWMLIAFAHGDLYAAMEKGTDLADPYVDVDPVYAGEGTVNPCVTSIHSSSAFLF 1680
 Qy 1681 SIEVOVTIGFGGMWTEECPLATILLIIVQNVGLMNTAMLGCI PMKTAQAHRAETLIF 1740
 Db 1681 SIEVOVTIGFGGMWTEECPLATILLIIVQNVGLMNTAMLGCI PMKTAQAHRAETLIF 1740
 Qy 1741 SKRAVITLRRGRCLCFMLRVGDLRKSMIISATIHQVVRKTSPEGEVPLHQVDIPMENG 1800
 Db 1741 SKRAVITLRRGRCLCFMLRVGDLRKSMIISATIHQVVRKTSPEGEVPLHQVDIPMENG 1800
 Qy 1801 VGGNGIFLVAPLLIIYHVIDSNSPLVDLAPSDLHHQDLEIIVILEGVETTGITTOARTS 1860
 Db 1801 VGGNGIFLVAPLLIIYHVIDSNSPLVDLAPSDLHHQDLEIIVILEGVETTGITTOARTS 1860
 Qy 1861 YLADELIMGQRFPIYAEEDGRYSVDYKRGNTIKVPTPCTAROLDSDSLDALTLAS 1920
 Db 1861 YLADELIMGQRFPIYAEEDGRYSVDYKRGNTIKVPTPCTAROLDSDSLDALTLAS 1920
 Qy 1921 SRGPLRKRSAVAVAKAPKFSISPDLS 1947
 Db 1901 SRGPLRAASVAVAKAPKFSISPDLS 1927

RESULT 4
 ADY86941
 ID ADY86941 standard; protein; 1891 AA.
 XX
 AC ADY86941;
 XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Human MRP1-Kir6.2 mutant delctac36 fusion protein, SRQ ID NO: 4.
 XX
 KM Ionophore; biosensor; drug screening; diagnostic;
 KM microorganism detection; potassium channel; fusion protein;
 KM multidrug resistance protein 1; MRP1; Kir6.2; mutcin.
 OS Homo sapiens.
 OS Mus musculus.
 OS Chimeric.
 XX
 FH Key
 FT Region 1..1531
 FT /note= "Multidrug resistance protein 1 (MRP1)"
 FT Region 1532..1537
 FT /note= "Hexaglycine spacer"
 FT Region 1538..1891
 FT /note= "Kir6.2 mutant delctac36 protein"
 PN US2005063989-A1.
 XX
 PD 24-MAR-2005.
 XX
 PF 22-SEP-2003; 2003US-00665283.
 XX
 PR 22-SEP-2003; 2003US-00665283.
 XX
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 XX
 PI Derand R, Garcia E, Prost A, Revilloud J, Viaudou M;
 XX
 DR WPI; 2005-252611/26.

XX New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
PS Claim 25; SEQ ID NO 4; 78pp; English.
XX
CC The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 mutant detac36 protein.
XX
SQ Sequence 1891 AA;
Query Match 96.9%; Score 9704; DB 9; Length 1891;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 20; Gaps 2;
QY 1 MALRGFSADGSDPLMDMNTWNTSNPDFTKCFQNTLVWVPCFYLMACFPFYLYLSRH 60
DB 1 MALRGFSADGSDPLMDMNTWNTSNPDFTKCFQNTLVWVPCFYLMACFPFYLYLSRH 60
QY DRGYIQMTPLNKTCTALGFLIMIVCMADLFYSPMERSGIFLAVPLVSPILGITTLLA 120
DB DRGYIQMTPLNKTCTALGFLIMIVCMADLFYSPMERSGIFLAVPLVSPILGITTLLA 120
QY 61 DRGYIQMTPLNKTCTALGFLIMIVCMADLFYSPMERSGIFLAVPLVSPILGITTLLA 120
DB 61 DRGYIQMTPLNKTCTALGFLIMIVCMADLFYSPMERSGIFLAVPLVSPILGITTLLA 120
QY 121 TPLIQLERRKGVSGSGLTFMWLVAVCALAIRSKIMTALKEDAOYDLFEDITFYVYS 180
DB 121 TPLIQLERRKGVSGSGLTFMWLVAVCALAIRSKIMTALKEDAOYDLFEDITFYVYS 180
QY 181 LLLIQVLSCGSDPSPLFSETIHDNCPRESSAFLSRITFWMTTGLIVGRPLGSD 240
DB 181 LLLIQVLSCGSDPSPLFSETIHDNCPRESSAFLSRITFWMTTGLIVGRPLGSD 240
QY 241 LMSLNKEDTSEQVPLVVKMKKECAKTRKOPVYVYSSKDPAOPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVPLVVKMKKECAKTRKOPVYVYSSKDPAOPKSSKVDANEVEAL 300
QY 301 IVSPQKEMNPSEFLKVLKYTFGPFLMSFPFKAIHDLMPGSGQILKLKIFVNDTRAP 360
DB 301 IVSPQKEMNPSEFLKVLKYTFGPFLMSFPFKAIHDLMPGSGQILKLKIFVNDTRAP 360
QY 361 WQGYFTYVLLFTVACLOTVLVHOYFHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
DB 361 WQGYFTYVLLFTVACLOTVLVHOYFHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVVAORFMDLATYINMISAPLOVILALYLLMLNGLPSVLAVGVAWVLAMPVN 480
DB 421 GEIYNLMSVVAORFMDLATYINMISAPLOVILALYLLMLNGLPSVLAVGVAWVLAMPVN 480
QY 481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYKUYAMELAFKDKVLAIROBELKVLK 540
DB 481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYKUYAMELAFKDKVLAIROBELKVLK 540
QY 541 KSAVLAVGFTTWCTPPLVALCTFPAVYVITDENNIIDAQTAFLSLFNLIRFPLNLP 600
DB 541 KSAVLAVGFTTWCTPPLVALCTFPAVYVITDENNIIDAQTAFLSLFNLIRFPLNLP 600
QY 601 MVISSIVQASVSLKRLRILPLSHEELPDSIERRPVKGGGNSITVNAATTAARSDPT 660
DB 601 MVISSIVQASVSLKRLRILPLSHEELPDSIERRPVKGGGNSITVNAATTAARSDPT 660
QY 661 LMGITFSPGALVAVVGVGCGKSSLSALLAEMDKVEGVAIKGSVAVYPQQAQND 720
DB 661 LMGITFSPGALVAVVGVGCGKSSLSALLAEMDKVEGVAIKGSVAVYPQQAQND 720
QY 721 SLRENIIFGCOLLEPPYRSVTOACALLPDLIELPSGDRTEIGEGVNLSSGQKORVSLAR 780
DB 721 SLRENIIFGCOLLEPPYRSVTOACALLPDLIELPSGDRTEIGEGVNLSSGQKORVSLAR 780

DB 721 SLRENIIFGCOLLEPPYRSVTOACALLPDLIELPSGDRTEIGEGVNLSSGQKORVSLAR 780
QY 781 AVYSNADIVLEDDPLSAVDAAVGHKIFENVIGRGMKNKTRILVTHSMYLPQVDVIY 840
DB 781 AVYSNADIVLEDDPLSAVDAAVGHKIFENVIGRGMKNKTRILVTHSMYLPQVDVIY 840
QY 841 MSGGKISEMGSYOBLLARDGAFAPLRTYASTQEOBDAENGVTVSGPGKEAKOMENG 900
DB 841 MSGGKISEMGSYOBLLARDGAFAPLRTYASTQEOBDAENGVTVSGPGKEAKOMENG 900
QY 901 LVNDSAGKOLQORLSSSSSYSGDISRHNSHTALQRAKKEFTWKLMEADKQOTGVKL 960
DB 901 LVNDSAGKOLQORLSSSSSYSGDISRHNSHTALQRAKKEFTWKLMEADKQOTGVKL 960
QY 961 SYVWDYKAIQLEIFSLFSLFPMCNHVSALASNYMLSLMTDPIVNGTOHTKVRLSVYG 1020
DB 961 SYVWDYKAIQLEIFSLFSLFPMCNHVSALASNYMLSLMTDPIVNGTOHTKVRLSVYG 1020
QY 1021 ALGISGIAVFGYSMAVSIIGIILASRCLHYDLHSILRSFMSFPERTSGNLVNRSEKEL 1080
DB 1021 ALGISGIAVFGYSMAVSIIGIILASRCLHYDLHSILRSFMSFPERTSGNLVNRSEKEL 1080
QY 1081 DTVDGMTPEVIKMFMSLFEVIGACIVILATPIAIIIPPLGLIYFVQRFVYASSROL 1140
DB 1081 DTVDGMTPEVIKMFMSLFEVIGACIVILATPIAIIIPPLGLIYFVQRFVYASSROL 1140
QY 1141 KRLESVRSFVYSHFNELLGVSIVIRAFEBQERFIHQSDLKVDENQAYYPSIVANRWLA 1200
DB 1141 KRLESVRSFVYSHFNELLGVSIVIRAFEBQERFIHQSDLKVDENQAYYPSIVANRWLA 1200
QY 1201 VRLECGNCIVLPAALFAVISRHSLSAGLVLSVSLQVTTYLNLVRRMSEMETNIVA 1260
DB 1201 VRLECGNCIVLPAALFAVISRHSLSAGLVLSVSLQVTTYLNLVRRMSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFRNYCLARYEDDLFVLRHINVTNGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFRNYCLARYEDDLFVLRHINVTNGG 1320
QY 1321 EKVGIVRTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGIIRMLNLPFSQYSDERVTSLELAHLKDFVSAIPKLDHBCAEGGENTSVGQROLVCL 1440
DB 1381 SGIIRMLNLPFSQYSDERVTSLELAHLKDFVSAIPKLDHBCAEGGENTSVGQROLVCL 1440
QY 1441 ARLALRRTKTLVDEATAVADLETDDLIOSTIRFOEDCTVLTAHRLNTIMOTRYIVL 1500
DB 1441 ARLALRRTKTLVDEATAVADLETDDLIOSTIRFOEDCTVLTAHRLNTIMOTRYIVL 1500
QY 1501 DKGEIOEGAPSDILOQRGLFYSMADKAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
DB 1501 DKGEIOEGAPSDILOQRGLFYSMADKAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
QY 1561 PRYTRERRARFVSKKNCNCVAHKNIREQGRPLQDVFTTLVDLKMPHTLLIFTMSFLCSW 1620
DB 1561 PRYTRERRARFVSKKNCNCVAHKNIREQGRPLQDVFTTLVDLKMPHTLLIFTMSFLCSW 1620
QY 1621 LLPAMVWMLLAPAHGDIYAMKEGIDPLAEPYDPVRYABEGTVPCVTSIHSSSAFLF 1680
DB 1621 LLPAMVWMLLAPAHGDIYAMKEGIDPLAEPYDPVRYABEGTVPCVTSIHSSSAFLF 1680
QY 1681 SIEVQVITFGGNNVTECEPLAIIILVONIYGLMNAIMLGCIFMTCTAQAHRAEFLIF 1740
DB 1681 SIEVQVITFGGNNVTECEPLAIIILVONIYGLMNAIMLGCIFMTCTAQAHRAEFLIF 1740
QY 1741 SKAVITLIRHGRCLCFMLRVGDLKSMIISATIHQVVRKTSPEGEVPLPHQVDIPMENG 1800
DB 1741 SKAVITLIRHGRCLCFMLRVGDLKSMIISATIHQVVRKTSPEGEVPLPHQVDIPMENG 1800
QY 1801 VGGNGITLVAPLIYHVIDNSPLXYDLABSDLHHQDLLEITVILEGIVETTGITTOARTS 1860
DB 1801 VGGNGITLVAPLIYHVIDNSPLXYDLABSDLHHQDLLEITVILEGIVETTGITTOARTS 1860
QY 1861 VGGNGITLVAPLIYHVIDNSPLXYDLABSDLHHQDLLEITVILEGIVETTGITTOARTS 1920
DB 1861 VGGNGITLVAPLIYHVIDNSPLXYDLABSDLHHQDLLEITVILEGIVETTGITTOARTS 1920

QY 1861 YLADEILMGORFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTAROLDEDRS 1911
 DB 1841 YLADEILMGORFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTAROLDEDRS 1891

RESULT 5
 ID AAM57486
 AAM57486 standard; protein; 1531 AA.
 AAM57486;
 AAM57486;
 14-AUG-1998 (first entry)
 Human MRP variant 1tpgpa (lei/pgpa).
 Multidrug resistance-associated protein; MRP; tumour; human; variant;
 multidrug resistance; MDR; leishmania P-glycoprotein; 1tpgpa, lei/pgpa.
 Homo sapiens.
 OS XX
 FH Key Location/Qualifiers
 FT Misc-difference 685 /label= L685S
 FT /note= "wild-type Leu is replaced by Ser"
 FT Misc-difference 1282 /label= R1282A
 FT /note= "wild-type Arg is replaced by Ala"
 FT US5766880-A.
 PN 16-JUN-1998.
 PD 05-JUN-1995; 95US-00463092.
 PF 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX (TOOH) UNIV QUEENS KINGSTON.
 PA Cole SP, Deeley RG;
 PI WPI; 1998-361687/31.
 DR N-PSDB; AAV31498.
 DR DNA encoding protein associated with multi-drug resistance - useful for
 PT as probe for identifying multi-drug resistant tumour cells.
 XX Claim 1; Col 67-78; 82pp; English.
 CC This represents a variant of the human multidrug resistance-associated
 CC protein (MRP). This natural variant is a leishmania P-glycoprotein related
 CC molecule 1tpgpa (lei/pgpa). The human and murine MRP nucleic acid
 CC molecules can be used as probes for identifying multidrug resistant
 CC tumour cells by hybridisation to mRNA from tumour cells. The antisense
 CC nucleic acid can be used to reverse multidrug resistance (MDR). A
 CC recombinant expression vector containing the MRP nucleic acid molecules
 CC operatively linked to at least one regulatory sequence can be used to
 CC transform a host cell to produce a recombinant MDR-associated protein
 XX Sequence 1531 AA;
 SO

Query Match 78.5%; Score 7860; DB 2; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARGPESAGSPBIMDMNTNTSNDFTKCFONTLVVWPFCYLLMACFPFFLYLSRH 60
 DB 1 MARGPESAGSPBIMDMNTNTSNDFTKCFONTLVVWPFCYLLMACFPFFLYLSRH 60
 QY 61 DRGTYQWTPINKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPFLIGITTLA 120

DB 61 DRGTYQWTPINKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPFLIGITTLA 120
 QY 121 TPLIQERRRGVSSGIMLTFWLVALVCAIAIIRSKITMTALKEDAOVDLPDITFYVYS 180
 DB 121 TPLIQERRRGVSSGIMLTFWLVALVCAIAIIRSKITMTALKEDAOVDLPDITFYVYS 180
 QY 181 LLLIQVLSCPSRSPLFSETHDNPCESSASFSLRIFPMWTTGLIVGYRQPLGSD 240
 DB 181 LLLIQVLSCPSRSPLFSETHDNPCESSASFSLRIFPMWTTGLIVGYRQPLGSD 240
 QY 241 LMSLNKEDTSEQVVPVLVKNWKECACTRRQPVVYSSKDPAPQKSSKVDAAEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKNWKECACTRRQPVVYSSKDPAPQKSSKVDAAEVEAL 300
 QY 301 IVKSPQKSNPSLFKVLKTFPGFYFLMSFPFKAIHDMFSGPQILKLIFVNDYAPD 360
 DB 301 IVKSPQKSNPSLFKVLKTFPGFYFLMSFPFKAIHDMFSGPQILKLIFVNDYAPD 360
 QY 361 MGGFYTVLLFVTAQTLVHOVHTCFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
 DB 361 MGGFYTVLLFVTAQTLVHOVHTCFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
 QY 421 GEIVNLMSVDAOREFMDLATYINNIWSAPLQVITLALYLLMTNLGPSVLAGVAVMVLAMYVN 480
 DB 421 GEIVNLMSVDAOREFMDLATYINNIWSAPLQVITLALYLLMTNLGPSVLAGVAVMVLAMYVN 480
 QY 481 AVAMKTKTYQVAMHMSKDNRIKLMEIINGIKVLKYLAMELAFKDYLAIROBELKVLK 540
 DB 481 AVAMKTKTYQVAMHMSKDNRIKLMEIINGIKVLKYLAMELAFKDYLAIROBELKVLK 540
 QY 541 KSAVLSAVGTFPTWVCTPFLVALCTFAVYVITIDENIIDAQTAFLVALFNIILRPILILP 600
 DB 541 KSAVLSAVGTFPTWVCTPFLVALCTFAVYVITIDENIIDAQTAFLVALFNIILRPILILP 600
 QY 601 MVISSIVQASVSLKRLRIFLSHELEBPSIERRPVKDGGSNTSYVANAFTMARSDPPT 660
 DB 601 MVISSIVQASVSLKRLRIFLSHELEBPSIERRPVKDGGSNTSYVANAFTMARSDPPT 660
 QY 661 LMGITFSPGALVAVVGQVCGKSSLSALLAEMDVKEGVAIKGSVAVVPQAWIQND 720
 DB 661 LMGITFSPGALVAVVGQVCGKSSLSALLAEMDVKEGVAIKGSVAVVPQAWIQND 720
 QY 721 SLRENIIFGQLEBPYRSYIQACALLPDLEILPBGDRTEIGKGVNLSCGOKORVSLAR 780
 DB 721 SLRENIIFGQLEBPYRSYIQACALLPDLEILPBGDRTEIGKGVNLSCGOKORVSLAR 780
 QY 781 AVYSNADIVLFDDPLSAVDHVGKHLFENVYIGPKMLKNTKRLVLYHMSGYLPOVDVITY 840
 DB 781 AVYSNADIVLFDDPLSAVDHVGKHLFENVYIGPKMLKNTKRLVLYHMSGYLPOVDVITY 840
 QY 841 MSGGKISEMGSYOELLARDGAPAEFLRTYASTQEDDAENGVGVSGPGEKAKOMENG 900
 DB 841 MSGGKISEMGSYOELLARDGAPAEFLRTYASTQEDDAENGVGVSGPGEKAKOMENG 900
 QY 901 LVTDSAGKQLOROLSSSSSYSGDISRHNSTALQYRAKKEETWKLMEADKQYQVYL 960
 DB 901 LVTDSAGKQLOROLSSSSSYSGDISRHNSTALQYRAKKEETWKLMEADKQYQVYL 960
 QY 961 SYVMDYMKAIGLIFSLIFLFMCNHYVSAASNYWYLSIMTDDPIVNGTOHTRVRLSYVG 1020
 DB 961 SYVMDYMKAIGLIFSLIFLFMCNHYVSAASNYWYLSIMTDDPIVNGTOHTRVRLSYVG 1020
 QY 1021 ALGISQIAVFGYSMAVSIIGILIASRCLVLDLHSLIRSPMSFERTPSGNLVNRSKEL 1080
 DB 1021 ALGISQIAVFGYSMAVSIIGILIASRCLVLDLHSLIRSPMSFERTPSGNLVNRSKEL 1080
 QY 1081 DTVDSMIPVYIKRPMGSLPRVIGACTVILLATPIAIIIPPLGLIYFFVQRFVYASSROL 1140
 DB 1081 DTVDSMIPVYIKRPMGSLPRVIGACTVILLATPIAIIIPPLGLIYFFVQRFVYASSROL 1140
 QY 1141 KRLESVRSFVYSHFNETLLGVSVIRAFEOEFHQSQDKVDENKAYPSIVANRWLA 1200

Db 1141 KRLESVSRSPVYSHFNETLIGVSVIRAFEEQERFIHOSDLKVDENOKAYPYSIVANRWLA 1200
QY 1201 VRLECVNGNCIVLFAALPAVISRHSLSAGLVGLSVYSIQVTTYINMLVRMSSEMETNIVA 1260
Db 1201 VRLECVNGNCIVLFAALPAVISRHSLSAGLVGLSVYSIQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSERTEKAPWQIQETAPPSWQVGRVEFRNYCLARYEDLDVLRHINVTINGG 1320
Db 1261 VERLKEYSERTEKAPWQIQETAPPSWQVGRVEFRNYCLARYEDLDVLRHINVTINGG 1320
QY 1321 EKVGI VERTGAGKSLTGLFRINESAGEIITIDGINAKIGHDLPKRTITIDDPVLF 1380
Db 1321 EKVGI VERTGAGKSLTGLFRINESAGEIITIDGINAKIGHDLPKRTITIDDPVLF 1380
QY 1381 SGSLRMNLDPSPQYSDSEWMTSLBLAHLDKDFVSLPDLDEHCAEGENLSVGOROLVCL 1440
Db 1381 SGSLRMNLDPSPQYSDSEWMTSLBLAHLDKDFVSLPDLDEHCAEGENLSVGOROLVCL 1440
QY 1441 APALLRRTKILVDEBATAVDLETDLDLIQSTIRFOEDCTVLTAAHRLNTIMDYTRIVL 1500
Db 1441 APALLRRTKILVDEBATAVDLETDLDLIQSTIRFOEDCTVLTAAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGAAPSDLLQORGLFYMAKXAGLV 1531
Db 1501 DKGEIOEYGAAPSDLLQORGLFYMAKXAGLV 1531

RESULT 6

AAM74471

ID AAM74471 standard; protein; 1531 AA.

AC AAM74471;

DT 18-MAY-1999 (first entry)

DE Human multidrug resistance-associated protein variant.

KM Multidrug resistance-associated protein; MDR; human; diagnosis;
MDR tumour cell identification; cancer therapy.

OS Homo sapiens.

PN US5882875-A.

PD 16-MAR-1999.

PE 05-JUN-1995; 95US-00462109.

PR 27-OCT-1992; 92US-00966923.

PR 08-MAR-1993; 93US-00029340.

PR 26-OCT-1993; 93US-00141893.

PR 20-MAR-1995; 95US-00407207.

PA (TOOH) UNIV QUEBENS KINGSTON.

PI Cole SPC, Deeley RG;

DR WPI: 1999-214061/18.

PT N-PSDB; AAX21977.

PT Identifying a multidrug resistant tumor cell by contacting the cell with
an antibody/antigen-binding fragment - which binds to an expressed
protein encoded by multidrug resistance-associated protein (MRP) nucleic
acid.

PS Claim 3; Col 69-80; 80pp; English.

CC This sequence is the human multidrug resistance-associated (MDR) protein.
CC The invention relates to a method for identifying a multidrug resistant
CC (MDR) tumor cell. Compositions and methods utilizing the MDR proteins
CC can be used to treat patients with tumours displaying multidrug
CC resistance, particularly those displaying resistance to anthracyclines,
CC epipodophyllotoxins, vinca alkaloids, and hydrophobic drugs. The methods

CC for inhibiting/killing a MDR tumour cell can be useful for treating
CC breast cancer, leukemias, fibrosarcomas, cervical cancer, gliomas,
CC thymomas, neuroblastomas and lung cancer. The MDR DNA sequences when
CC labeled are useful as molecular probes for diagnosing multidrug
CC resistance of a tumour (using cells from a tumour biopsy) and for
CC designing ribozymes which are capable of cleaving a single-stranded
CC nucleic acid encoding a protein having MRP activity. Recombinant
CC expression vectors containing human MDR coding sequences can be
CC transfected into a drug sensitive cell line to produce a protein in the
CC cell which confers MDR, protecting non-resistant non-tumour cells from
CC the effects of chemotherapeutics has major clinical importance. Cells
CC transformed with the MDR coding sequences are useful for testing
CC potential therapeutic agents for their effectiveness against MDR cells
CC and for identifying chemosensitizers of a therapeutic agent

XX Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVMTNSNDFTKCFQNTVLVWVPCPYLMACFPFYLYSRH 60
Db 1 MALRGFCSADGSDPLMDMNTVMTNSNDFTKCFQNTVLVWVPCPYLMACFPFYLYSRH 60
QY 61 DRGYIQMTPLNKTKTALGFLIMIVCWADLFYSPWERSRGIFLAEVFLVSPITLIGITLLA 120
Db 61 DRGYIQMTPLNKTKTALGFLIMIVCWADLFYSPWERSRGIFLAEVFLVSPITLIGITLLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFLMLVYALCALAIRSKMTALKEBAQVDLPDRTIFYFYFS 180
Db 121 TFLIQLERRKGVQSSGIMLTFLMLVYALCALAIRSKMTALKEBAQVDLPDRTIFYFYFS 180
QY 181 LLLIQVLVSCPSDPSPLFSETIHDNPNCPSSASFLSRTIFPMWITGLIVRGYROPLEGSD 240
Db 181 LLLIQVLVSCPSDPSPLFSETIHDNPNCPSSASFLSRTIFPMWITGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVYSSXDPAPKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVYSSXDPAPKSSKYDANEVEAL 300
QY 301 IVKSPQKBNPNSLFKVLKYTRGPYFLMSFPFKAIHDLMSFGPOLIKLLIKFVNDYAPD 360
Db 301 IVKSPQKBNPNSLFKVLKYTRGPYFLMSFPFKAIHDLMSFGPOLIKLLIKFVNDYAPD 360
QY 361 WQGFYTVLLEFVTACLOTLVHGYEHI CFVSGMKIKTAVGAVYRKALVITNSARKSTV 420
Db 361 WQGFYTVLLEFVTACLOTLVHGYEHI CFVSGMKIKTAVGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQREMDLATYINMTWSAPLQVITLALYLLMNLGFSVLGAVAVWVLMVNVN 480
Db 421 GEIVNLMSVDAQREMDLATYINMTWSAPLQVITLALYLLMNLGFSVLGAVAVWVLMVNVN 480
QY 481 AVMAKTKTYOVAMKSKDNRIKLMNEILNGIKYIKIYAMELAKDKYLAIRROELKYLK 540
Db 481 AVMAKTKTYOVAMKSKDNRIKLMNEILNGIKYIKIYAMELAKDKYLAIRROELKYLK 540
QY 541 KSAVLASVGFETWCTCFVVALCTFAVYVITDENNNIIIDAOFAFSLAFNLRPEPLNLP 600
Db 541 KSAVLASVGFETWCTCFVVALCTFAVYVITDENNNIIIDAOFAFSLAFNLRPEPLNLP 600
QY 601 MVISSIVQASVSLKRLNIFLSHELEPDSIERPVKDGGSITVNVNATPTWARSDEPT 660
Db 601 MVISSIVQASVSLKRLNIFLSHELEPDSIERPVKDGGSITVNVNATPTWARSDEPT 660
QY 661 LNGITFSIPREGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAMTQND 720
Db 661 LNGITFSIPREGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAMTQND 720
QY 721 SLRENIILFGCLLEPPYRSVIAQCALLPDLEIILSGRTEIGEGNVNLSGGOKRVSLAR 780
Db 721 SLRENIILFGCLLEPPYRSVIAQCALLPDLEIILSGRTEIGEGNVNLSGGOKRVSLAR 780

QY 781 AVYSNADLYLFDDEPLSAVDHVGKHIIFENVIGPKMKLKNKTRILVTHSMSTLPQVDYIIV 840
 DB 781 AVYSNADLYLFDDEPLSAVDHVGKHIIFENVIGPKMKLKNKTRILVTHSMSTLPQVDYIIV 840
 QY 841 MSGGKISEMGSYOELARDGAFALFNTYASTEOBDAEENGVTGSGPKKAYOMENGM 900
 DB 841 MSGGKISEMGSYOELARDGAFALFNTYASTEOBDAEENGVTGSGPKKAYOMENGM 900
 QY 901 LVTDSACKOLOROLSSSSSSSGDISRRHNSTAEIOKAEAKEEFTWKLMEADKATQGVKL 960
 DB 901 LVTDSACKOLOROLSSSSSSSGDISRRHNSTAEIOKAEAKEEFTWKLMEADKATQGVKL 960
 QY 961 SVYWDYKKAIGLFIISFISIFLFCMCHVSAALSNYMLSLMTDPIVNGTOBHTKRLSVYG 1020
 DB 961 SVYWDYKKAIGLFIISFISIFLFCMCHVSAALSNYMLSLMTDPIVNGTOBHTKRLSVYG 1020
 QY 1021 ALGISOGIAVGVGMAVSIIGGILASRCLHVDLHSHLSRSPMSFEPTPSGNLVNRFSEKL 1080
 DB 1021 ALGISOGIAVGVGMAVSIIGGILASRCLHVDLHSHLSRSPMSFEPTPSGNLVNRFSEKL 1080
 QY 1081 DTWDSMTPEYIKYKMGSLFNVIAGCIYITLAPTAIITIPPLGLIYFFVQAFYASSROL 1140
 DB 1081 DTWDSMTPEYIKYKMGSLFNVIAGCIYITLAPTAIITIPPLGLIYFFVQAFYASSROL 1140
 QY 1141 KRLESVSRSPVSHFNFTLGVSVYIRAFEBEORFIHOSDLKVDENOKAYPSIVANRMLA 1200
 DB 1141 KRLESVSRSPVSHFNFTLGVSVYIRAFEBEORFIHOSDLKVDENOKAYPSIVANRMLA 1200
 QY 1201 VRLECVNGCIYLFALPAVIRSRHLSAGLVGLSVSYSLQVTTYLVNMLVRMSSEMETIVA 1260
 DB 1201 VRLECVNGCIYLFALPAVIRSRHLSAGLVGLSVSYSLQVTTYLVNMLVRMSSEMETIVA 1260
 QY 1261 VERIKESSETEKAPMOIOETAPSSWPQVGRVFRNYCIRREDDLPVLRHINVTINGG 1320
 DB 1261 VERIKESSETEKAPMOIOETAPSSWPQVGRVFRNYCIRREDDLPVLRHINVTINGG 1320
 QY 1321 EKYGIIVGRTAGSGSLTLGLFRINESAGEIITIDGINIAKIGLHDLFKITIIPODVLV 1380
 DB 1321 EKYGIIVGRTAGSGSLTLGLFRINESAGEIITIDGINIAKIGLHDLFKITIIPODVLV 1380
 QY 1381 SSGSLRNMLDPPSQYSDSEEVWTSLELAHLKDFVSALPKLDHECAEGENISVGRQLVCL 1440
 DB 1381 SSGSLRNMLDPPSQYSDSEEVWTSLELAHLKDFVSALPKLDHECAEGENISVGRQLVCL 1440
 QY 1441 ARALLRRTKTLVDEATAVADLETDLIQSTIRTOPEBDCTVLITAHRLNTIMDYTRIVL 1500
 DB 1441 ARALLRRTKTLVDEATAVADLETDLIQSTIRTOPEBDCTVLITAHRLNTIMDYTRIVL 1500
 QY 1501 DKGEIOEYGA PSDLLOQRGLFYSMAKDAGLV 1531
 DB 1501 DKGEIOEYGA PSDLLOQRGLFYSMAKDAGLV 1531

RESULT 7
 AAM99894
 ID AAM99894 standard; protein; 1531 AA.
 AC AAM99894;
 DT 10-JUN-1999 (first entry)
 XX
 DE Human multidrug resistance-associated protein natural variant.
 KW Human; multidrug resistance-associated protein; MRP; cytotoxic drug;
 KW cancer; chemotherapy.
 OS Homo sapiens.
 OS
 PN US5891724-A.
 XX
 XX 06-APR-1999.
 PD
 PF 05-JUN-1995; 95US-00460907.

XX 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX
 PA (TOOH) UNIV QUEBENS KINGSTON.
 XX
 PI Cole SPC, Deeley RG;
 XX
 XX WPI; 1999-253868/21.
 DR N-PSDB; AAX19818.
 DR
 XX
 PT Protecting mammalian cells against cytotoxic drugs.
 XX
 PS Claim 2; Col 79-86; 82pp; English.
 CC
 CC The present sequence represents a human multidrug resistance-associated
 CC protein (MRP). The present invention also describes a method for
 CC protecting a mammalian cell against the cytotoxicity of anthracyclines,
 CC epipodophyllotoxins and Vinca alkaloids (A) by introducing into it a
 CC nucleic acid (I) that hybridizes under stringent conditions to a nucleic
 CC acid (II) that encodes an MRP protein (III). Introduction of (I) protects
 CC cells against cytotoxic effects of (A), particularly to protect normal
 CC cells against (A) being used for treatment of cancers. Cells transformed
 CC with (I) can be used to screen for agents that affect multidrug
 CC resistance or are directly toxic to multidrug resistant cells, i.e.
 CC potential therapeutics for multidrug-resistant cancers. Confering
 CC resistance to normal cells should allow an increase in the dose of (A)
 CC that can be administered safely
 XX
 SQ Sequence 1531 AA:
 Query Match 78.5%; Score 7860; DB 2; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MALRGFCSADGSDPLMDNVTNTSNPDKTCQNTYLVWVPCFYLMACPPFFLYLSRH 60
 DB 1 MALRGFCSADGSDPLMDNVTNTSNPDKTCQNTYLVWVPCFYLMACPPFFLYLSRH 60
 QY 61 DRGYIQMTPLNKTKTLGLFLLWVCAADLFSWERSRGFLAPVFLVSTLIGITTLA 120
 DB 61 DRGYIQMTPLNKTKTLGLFLLWVCAADLFSWERSRGFLAPVFLVSTLIGITTLA 120
 QY 121 TFLIOERRKGVSSGIMLTFMLVALVICALALIRSKIMTALKEADAQVDLFRDITFYVFS 180
 DB 121 TFLIOERRKGVSSGIMLTFMLVALVICALALIRSKIMTALKEADAQVDLFRDITFYVFS 180
 QY 181 LLLIOVLSCFSGRSPLFSETTHDNPCESSASFSRITFWITGLIYRGYQPLBGSD 240
 DB 181 LLLIOVLSCFSGRSPLFSETTHDNPCESSASFSRITFWITGLIYRGYQPLBGSD 240
 QY 241 LWSLNKEDPSEGVVPLVKNWKECAKTRQPKVYSSDPOPKSSSVVDNNEVEAL 300
 DB 241 LWSLNKEDPSEGVVPLVKNWKECAKTRQPKVYSSDPOPKSSSVVDNNEVEAL 300
 QY 301 IVKSPQKWNPSLFLKVLKTFGPFYLMSPFFKAIHDLMMFSGQIILKLIKFNNDTKAPD 360
 DB 301 IVKSPQKWNPSLFLKVLKTFGPFYLMSPFFKAIHDLMMFSGQIILKLIKFNNDTKAPD 360
 QY 361 WQGYFTVLLFVTAQCTVLVHOYFHI CPVSGMRITAVIGAYRRALVITNSARKSSTV 420
 DB 361 WQGYFTVLLFVTAQCTVLVHOYFHI CPVSGMRITAVIGAYRRALVITNSARKSSTV 420
 QY 421 GEIVNLSYDAQRFMDLATYINNIWASAPLOVITALLVLMNLGSPSYLAGAVVWLVWV 480
 DB 421 GEIVNLSYDAQRFMDLATYINNIWASAPLOVITALLVLMNLGSPSYLAGAVVWLVWV 480
 QY 481 AVAMKTKTYQVAHMSKONRIKLMEIINGIKVLKLYAMELAFKDKVLAIRQEBLKVK 540
 DB 481 AVAMKTKTYQVAHMSKONRIKLMEIINGIKVLKLYAMELAFKDKVLAIRQEBLKVK 540

```

QY 541 KSAVLSAVGFTFWCTPFLVALCTPAVYVITDENNIIDAQTAFLVALFNILRFPNLLP 600
DB 541 KSAVLSAVGFTFWCTPFLVALCTPAVYVITDENNIIDAQTAFLVALFNILRFPNLLP 600
QY 601 MVISSIVQASVSKRLRIPLSHELEPDSIRRPVKOGGNGNSTVNAATATMARSPPT 660
DB 601 MVISSIVQASVSKRLRIPLSHELEPDSIRRPVKOGGNGNSTVNAATATMARSPPT 660
QY 661 LNCITFSPREGALVAVVGVQVCGKSSLLSALLAEMDKVEGVAIKGSVAVYVPOQAMIOND 720
DB 661 LNCITFSPREGALVAVVGVQVCGKSSLLSALLAEMDKVEGVAIKGSVAVYVPOQAMIOND 720
QY 721 SLRENILFGCQLBEPYRSVIOACALLPDLEILPSGDRTEIGEKGVNLSGGQXORVSLAR 780
DB 721 SLRENILFGCQLBEPYRSVIOACALLPDLEILPSGDRTEIGEKGVNLSGGQXORVSLAR 780
QY 781 AVYNNADIYLFDDPLSAVDAAVGHIFENVIQPKGMLKNKRIILVTHSMSTLPQVDVIV 840
DB 781 AVYNNADIYLFDDPLSAVDAAVGHIFENVIQPKGMLKNKRIILVTHSMSTLPQVDVIV 840
QY 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEODAEENGVTGSGPGKEAKOMENGM 900
DB 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEODAEENGVTGSGPGKEAKOMENGM 900
QY 901 LVTDASAGKOLQROUSSSSSSSGDISRHHNSTAEIQAKAKKEETWKLMEADKAQTGYKL 960
DB 901 LVTDASAGKOLQROUSSSSSSSGDISRHHNSTAEIQAKAKKEETWKLMEADKAQTGYKL 960
QY 961 SVYVDYMKALIFLSLIFLPMCHVSAASNYWLSMTWDDPIVNGOETHKRLSYG 1020
DB 961 SVYVDYMKALIFLSLIFLPMCHVSAASNYWLSMTWDDPIVNGOETHKRLSYG 1020
QY 1021 ALGISOGIAVFGSYMAVSIIGGILASRCLHVDLHSLILSPMSFPERTSGNLVNRFSEL 1080
DB 1021 ALGISOGIAVFGSYMAVSIIGGILASRCLHVDLHSLILSPMSFPERTSGNLVNRFSEL 1080
QY 1081 DTVDSMIPEVIKMFMSLFENVIGACTIVILLATPIAIIIPGLGIYFFVQRPYVASSRQL 1140
DB 1081 DTVDSMIPEVIKMFMSLFENVIGACTIVILLATPIAIIIPGLGIYFFVQRPYVASSRQL 1140
QY 1141 KRLESVSRSPPYSHPHNETLIGSVITRAFEQERFPHQSDLVVDENQKAYYSIVANRWLA 1200
DB 1141 KRLESVSRSPPYSHPHNETLIGSVITRAFEQERFPHQSDLVVDENQKAYYSIVANRWLA 1200
QY 1201 VRLCEVNCIYLFALPAFVISRHSLSAGLVGASVYSIQVTTYLANWLVRMSEMETNIYA 1260
DB 1201 VRLCEVNCIYLFALPAFVISRHSLSAGLVGASVYSIQVTTYLANWLVRMSEMETNIYA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVEFRNYCLARYEDLDFVLRIHINTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVEFRNYCLARYEDLDFVLRIHINTINGG 1320
QY 1321 EKVGIVGRTGAKGSLTIGLFRINSAAGEIITIDGINIAKGLHDLRKITITIPDDPLF 1380
DB 1321 EKVGIVGRTGAKGSLTIGLFRINSAAGEIITIDGINIAKGLHDLRKITITIPDDPLF 1380
QY 1381 SGLSRMNLDPSPQSDSEEWMTSLIAHLKDFVSALPDKLDEHCAGGENLSVGQOLVCL 1440
DB 1381 SGLSRMNLDPSPQSDSEEWMTSLIAHLKDFVSALPDKLDEHCAGGENLSVGQOLVCL 1440
QY 1441 ARALLRKTKIIVLEBATAVVDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDYRIVVL 1500
DB 1441 ARALLRKTKIIVLEBATAVVDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDYRIVVL 1500
QY 1501 DKGEIOEYGAPSDLIQORGLFYSMAKAGLV 1531
DB 1501 DKGEIOEYGAPSDLIQORGLFYSMAKAGLV 1531

```

RESULT 8
AAVS5799 standard; protein; 1531 AA.
XX

```

AC AAV55799;
XX 28-FEB-2000 (first entry)
DT 28-FEB-2000 (first entry)
XX Human multidrug resistance-associated protein (MRP) variant.
DE Chemosensitizer; multidrug resistance-associated protein; MRP; human;
KW therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;
KW cancer; variant.
XX Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Key 685
FT Misc-difference /label= L685S
FT /note= "wild-type Leu is replaced by Ser"
FT Misc-difference 1282
FT /label= R1282A
FT /note= "wild-type Arg is replaced by Ala"
XX US6001563-A.
XX 14-DEC-1999.
XX 05-JUN-1995; 95US-00463179.
XX 27-OCT-1992; 92US-00966923.
XX 08-MAR-1993; 93US-00029340.
XX 26-OCT-1993; 93US-00141893.
XX 20-MAR-1995; 95US-00407207.
XX (TOOH ) UNIV QUEBENS KINGSTON.
XX Cole SP, Deeley RG;
XX WPI; 2000-061877/05.
XX DR N-PSDB; AAZ39556.
XX Identification of chemosensitizers useful for treating cancer, using
PT nucleic acid encoding multidrug resistance-associated protein.
XX Claim 3; Col 69-80; 77pp; English.
XX The invention provides a method for identifying a substance which is a
XX chemosensitizer that comprises, contacting a cell transfected with
XX nucleic acid encoding multidrug resistance-associated protein (MRP) with
XX a therapeutic agent in vitro. The method is useful for identifying
XX chemosensitizers which may then be used to treat cancer (especially lung
XX cancer). The method allows the identification of chemosensitizers which
XX do not reverse P-glycoprotein-mediated multidrug resistance. The present
XX sequence represents a human MRP variant
XX
SQ Sequence 1531 AA:
Query Match 78.5%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALMGFCAGDSDDLMMNMTVMTNSNDPTKCFONTYLVWPCRYLWACPPFYLYLSRH 60
DB 1 MALMGFCAGDSDDLMMNMTVMTNSNDPTKCFONTYLVWPCRYLWACPPFYLYLSRH 60
QY 61 DRGYIOMTPLNKKTALGFILMIYCMADLFYSFWEKSGTFLAVPLVSPFLIGITTLA 120
DB 61 DRGYIOMTPLNKKTALGFILMIYCMADLFYSFWEKSGTFLAVPLVSPFLIGITTLA 120
QY 121 TFLIQLEKRGVQSSGIMLTFWLVALVCALAILSKIMTALKEDAQVDLFRDITFYVYFS 180
DB 121 TFLIQLEKRGVQSSGIMLTFWLVALVCALAILSKIMTALKEDAQVDLFRDITFYVYFS 180
QY 181 LLLIQVLVSCPSDRSPFLSETIHDNPNCPBSSASFLSRTITWITGLIVRGYRQPLBGS 240
DB 181 LLLIQVLVSCPSDRSPFLSETIHDNPNCPBSSASFLSRTITWITGLIVRGYRQPLBGS 240

```


QY 241 LMSLNKEDTSEQVVPVLVKWKKECAKTRKOPVKVYSSKDPAPQKSSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKWKKECAKTRKOPVKVYSSKDPAPQKSSSKVDANEVEAL 300
 QY 301 IVKSPQKEMNPDLSEKVLKFTGPGYFLMSFFPKA1HDLMPFSGPOLIKLL1KFVNDTRAPD 360
 DB 301 IVKSPQKEMNPDLSEKVLKFTGPGYFLMSFFPKA1HDLMPFSGPOLIKLL1KFVNDTRAPD 360
 QY 361 WQGFYFVVLFPVTRCLOTLVHLQYFHI CFSVGNM1KTAIVGAVYRKALVITNSARKSTV 420
 DB 361 WQGFYFVVLFPVTRCLOTLVHLQYFHI CFSVGNM1KTAIVGAVYRKALVITNSARKSTV 420
 QY 421 GEIYNLMSVDAQRFMDLATYINMIWSA PLOYTLALYLLMLNGPSYAGVAVMWLAMPVN 480
 DB 421 GEIYNLMSVDAQRFMDLATYINMIWSA PLOYTLALYLLMLNGPSYAGVAVMWLAMPVN 480
 QY 481 AVNMAKTKTYQVAMKSKDNRIKLMNEILNGIKVLYAMELAFKDKVLA1ROBELKVLK 540
 DB 481 AVNMAKTKTYQVAMKSKDNRIKLMNEILNGIKVLYAMELAFKDKVLA1ROBELKVLK 540
 QY 541 KSAVLASVGTFTWCTPFLVALCTFPAVYVITDENNIIDAQTAFLVSLFNLIRPPLNLP 600
 DB 541 KSAVLASVGTFTWCTPFLVALCTFPAVYVITDENNIIDAQTAFLVSLFNLIRPPLNLP 600
 QY 601 MVISSIVQASVSLKRLIPLSHELEPDSIERRPVKDGGGNSITVNAFTTARSDPT 660
 DB 601 MVISSIVQASVSLKRLIPLSHELEPDSIERRPVKDGGGNSITVNAFTTARSDPT 660
 QY 661 LNCITFSP1PEGALVAVVGQVCGCKSSLSALLAEMDKYEGHVAIKGSVAVYVPOQAW1OND 720
 DB 661 LNCITFSP1PEGALVAVVGQVCGCKSSLSALLAEMDKYEGHVAIKGSVAVYVPOQAW1OND 720
 QY 721 SLRENILFGCOLBEPYRSYIQAICALIPDLLEILPSGRTEIGEKNVLSGGQKQVSLAR 780
 DB 721 SLRENILFGCOLBEPYRSYIQAICALIPDLLEILPSGRTEIGEKNVLSGGQKQVSLAR 780
 QY 781 AAVSNMDIYLPDPLSAVDAHVGHIFENYIGPQGM1KNTRIILVTHSMYLPQVDYIIV 840
 DB 781 AAVSNMDIYLPDPLSAVDAHVGHIFENYIGPQGM1KNTRIILVTHSMYLPQVDYIIV 840
 QY 841 MSGGKISEMSYOEILARDAFAEFLRTYASTEOQDAEENGVTGSGPGKEAKOMENG 900
 DB 841 MSGGKISEMSYOEILARDAFAEFLRTYASTEOQDAEENGVTGSGPGKEAKOMENG 900
 QY 901 LVTDGAKQULQROLSSSSSYSGDISRHNSIAELQKAEAKKETWKLMEADKAQTQVKL 960
 DB 901 LVTDGAKQULQROLSSSSSYSGDISRHNSIAELQKAEAKKETWKLMEADKAQTQVKL 960
 QY 961 SYVWDVYKATIGLFTSELIPLFMCNHVSALASNWLSLMTDDPIVNGTOEHTKRLSVYG 1020
 DB 961 SYVWDVYKATIGLFTSELIPLFMCNHVSALASNWLSLMTDDPIVNGTOEHTKRLSVYG 1020
 QY 1021 ALG1SGI1AVFGYSMAVISG1ILASRCLAHYDLHSILRSPMSFPERTPSGNLVNRFSEKL 1080
 DB 1021 ALG1SGI1AVFGYSMAVISG1ILASRCLAHYDLHSILRSPMSFPERTPSGNLVNRFSEKL 1080
 QY 1081 DTVDSMIPEVIKMFMSGLFNVI1GACIYILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
 DB 1081 DTVDSMIPEVIKMFMSGLFNVI1GACIYILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
 QY 1141 KRLIESVRSRPFVYGHFNFTLLGVSYTARFEEOERFIHOSDLKVDENQAAVYPSIVANWLA 1200
 DB 1141 KRLIESVRSRPFVYGHFNFTLLGVSYTARFEEOERFIHOSDLKVDENQAAVYPSIVANWLA 1200
 QY 1201 VRLIEVCNCTIVLPAALFAVISRSHLSAGVLGSYS1QVTTYLNMVLVRMSSEMETIYA 1260
 DB 1201 VRLIEVCNCTIVLPAALFAVISRSHLSAGVLGSYS1QVTTYLNMVLVRMSSEMETIYA 1260
 QY 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHINVTINGG 1320
 DB 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHINVTINGG 1320

QY 1321 EKVGIVRTGAGKSLTLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIITIPQDVLV 1380
 DB 1321 EKVGIVRTGAGKSLTLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIITIPQDVLV 1380
 QY 1381 SCSLRNMLDPPSOYSDEEVTSLSLAHLKDFVSALPDKLDHECAEGEENL SVGROLVCL 1440
 DB 1381 SCSLRNMLDPPSOYSDEEVTSLSLAHLKDFVSALPDKLDHECAEGEENL SVGROLVCL 1440
 QY 1441 ARAILRTKTLVDEATAAADLETDLIQSTITPOFEDCTVLT1AHLN1MTYTRIVYL 1500
 DB 1441 ARAILRTKTLVDEATAAADLETDLIQSTITPOFEDCTVLT1AHLN1MTYTRIVYL 1500
 QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKADAGLV 1531
 DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAKADAGLV 1531

RESULT 9
 ID AAY78873 standard; protein; 1531 AA.
 XX AAY78873;
 AC AAY78873;
 XX 19-MAY-2000 (first entry)
 DT 19-MAY-2000 (first entry)
 XX
 DE Multidrug resistance protein (MRP) natural variant amino acid sequence.
 XX
 KW Multidrug resistance protein; MRP; human; anthracycline; vinca alkaloid;
 KM epipodophyllotoxin; cancer; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN US6025473-A.
 XX 15-FEB-2000.
 PD 15-FEB-2000.
 XX
 PF 05-JUN-1995; 95US-00461384.
 XX
 PR 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX
 PA (TOOH) UNITV QUBENS KINGSTON.
 PA
 PI Cole SPC, Deeley RG;
 PI
 XX WPI; 2000-181838/16.
 DR N-PDB; AA290193.
 XX
 PT Isolated protein conferring multidrug resistance, to at least two drugs
 PT selected from anthracyclines, epipodophyllotoxins and vinca alkaloids, on
 PT a drug sensitive mammalian cell.
 XX
 PS Claim 10; Col 79-88; 78pp; English.
 PS
 XX This sequence represents a human multidrug resistance protein (MRP)
 XX natural variant amino acid sequence. The human MRP confers multidrug
 XX resistance, including resistance to at least two drugs selected from
 XX anthracyclines, epipodophyllotoxins and vinca alkaloids, on a drug
 XX sensitive mammalian cell, when the protein is expressed in the cell. The
 XX multidrug resistance is not substantially reversed by chemosensitizers
 XX which reverse P-glycoprotein-mediated multidrug resistance. The MRP
 XX protein sequence can be used to generate antibodies against MRP. The MRP
 XX protein and nucleotide sequences can be used in compositions which are
 XX used to treat patients with tumours displaying multidrug resistance. The
 XX compositions and methods of the invention can be used particularly to
 XX treat breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,
 XX thymomas, neuroblastomas, and lung cancer. Antibodies directed against
 XX MRP can be used to inhibit the multidrug resistance of a multidrug
 XX resistant cell
 XX
 XX Sequence 1531 AA;

Query Match 78.5%; Score 7860; Db 3; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFSSAGSDPLDMNTMTNTSNPDFTKCFONTVLVWPCTYLMACFPFYLYSRH 60
 Db 1 MALRGFSSAGSDPLDMNTMTNTSNPDFTKCFONTVLVWPCTYLMACFPFYLYSRH 60

QY 61 DRGIIQMTPLNKTALGFLIMIVCMADLFYFMSERSRGIFLAVVPLVSPILGITTLLA 120
 Db 61 DRGIIQMTPLNKTALGFLIMIVCMADLFYFMSERSRGIFLAVVPLVSPILGITTLLA 120

QY 121 TFLIQLERRKGVSSGIMLTFMWLVACALILRSKIMTAKEDAOYDLFFDITFYFFYS 180
 Db 121 TFLIQLERRKGVSSGIMLTFMWLVACALILRSKIMTAKEDAOYDLFFDITFYFFYS 180

QY 181 LLLIQLVLSCTSDSPSLFSETIHDNCPRESSASFSLRITFMWITGLIVRGYRQPLEGSD 240
 Db 181 LLLIQLVLSCTSDSPSLFSETIHDNCPRESSASFSLRITFMWITGLIVRGYRQPLEGSD 240

QY 241 LMSLNKEDTSBOVAVLVKMKKCAKTRKOPVAVVYSSKDPAPKSSKYDANEVEAL 300
 Db 241 LMSLNKEDTSBOVAVLVKMKKCAKTRKOPVAVVYSSKDPAPKSSKYDANEVEAL 300

QY 301 IVKSPQKMNPSLRFKVLKTFGPYFLMSFPFKA1HDLMPFSGPQILKLLIKFVNDTKAPD 360
 Db 301 IVKSPQKMNPSLRFKVLKTFGPYFLMSFPFKA1HDLMPFSGPQILKLLIKFVNDTKAPD 360

QY 361 WQGYFTYVLLFVTAQCLQTLVHQYFHCVSQMRK1KTAIVGAIVRKALVITNSARKSTV 420
 Db 361 WQGYFTYVLLFVTAQCLQTLVHQYFHCVSQMRK1KTAIVGAIVRKALVITNSARKSTV 420

QY 421 GEIYNLMSVDAORFMDLATYINMTWSAPLOVILALYLLMLGSPVLAVAWMLAMPVN 480
 Db 421 GEIYNLMSVDAORFMDLATYINMTWSAPLOVILALYLLMLGSPVLAVAWMLAMPVN 480

QY 481 AVNMAKTKTYOVAMKSKDNRIKLMN1LNGIKVLYAMELAFKDYLA1ROBELKVLK 540
 Db 481 AVNMAKTKTYOVAMKSKDNRIKLMN1LNGIKVLYAMELAFKDYLA1ROBELKVLK 540

QY 541 KSAVLSAVGFTTWCTPFLVALCTPAAYVTIDENN1LDAQTAFAVSLAFN1LRPPLN1LP 600
 Db 541 KSAVLSAVGFTTWCTPFLVALCTPAAYVTIDENN1LDAQTAFAVSLAFN1LRPPLN1LP 600

QY 601 MVISIVQASVSLKRLR1PLSHELEPDSIERRPVKOGGNS1TVNATFTWARSDEPT 660
 Db 601 MVISIVQASVSLKRLR1PLSHELEPDSIERRPVKOGGNS1TVNATFTWARSDEPT 660

QY 661 LMGITFS1PEGALVAVVQVCGKSS1LSALLAEMDKVEGHVA1KGSVAVYPQOAW1OND 720
 Db 661 LMGITFS1PEGALVAVVQVCGKSS1LSALLAEMDKVEGHVA1KGSVAVYPQOAW1OND 720

QY 721 SLREN11PGCOLLEPYRSV1OACALLPDL1LPSGDRTE1GEKGVN1SGGOKORVSLAR 780
 Db 721 SLREN11PGCOLLEPYRSV1OACALLPDL1LPSGDRTE1GEKGVN1SGGOKORVSLAR 780

QY 781 AVVSNAD1YLPDD1LSAVDAHVGH1PENNV1GPGKMLKNKRL1VITSM5MLPQVDV1IV 840
 Db 781 AVVSNAD1YLPDD1LSAVDAHVGH1PENNV1GPGKMLKNKRL1VITSM5MLPQVDV1IV 840

QY 841 MSGGK1SEMGSYQEL1LARDGAF1FLRTYASTEOQDAEENGVTGVS1PGKEAKQEMNGM 900
 Db 841 MSGGK1SEMGSYQEL1LARDGAF1FLRTYASTEOQDAEENGVTGVS1PGKEAKQEMNGM 900

QY 901 LVTD5AGKQ1ORQ1SSSSSSYS1SRHNS1TAEL1OKAARKEETWKLMEADKAOGTQVYL 960
 Db 901 LVTD5AGKQ1ORQ1SSSSSSYS1SRHNS1TAEL1OKAARKEETWKLMEADKAOGTQVYL 960

QY 961 SVVADYKAK1GL1F1S1FLFMCNHV1SALSNW1SLMTDDP1VNGQOHTKX1RLSYG 1020
 Db 961 SVVADYKAK1GL1F1S1FLFMCNHV1SALSNW1SLMTDDP1VNGQOHTKX1RLSYG 1020

QY 1021 ALGISG1AVFGYSMAV5IGG1IASRCLAVD1LH51IR5MSM5FEPTPSGNLVNRSKEL 1080
 Db 1021 ALGISG1AVFGYSMAV5IGG1IASRCLAVD1LH51IR5MSM5FEPTPSGNLVNRSKEL 1080

QY 1081 DTVD5MTPEV1K1KMFMS1ENV1GAC1V1LATP1A111PPLG1YFVORFVASSROL 1140
 Db 1081 DTVD5MTPEV1K1KMFMS1ENV1GAC1V1LATP1A111PPLG1YFVORFVASSROL 1140

QY 1141 KRLESVRS1PVY5HFN1ETL1G5V1RAFE5QE5F1HOSD1KVDENOKAYYPS1VANRWLA 1200
 Db 1141 KRLESVRS1PVY5HFN1ETL1G5V1RAFE5QE5F1HOSD1KVDENOKAYYPS1VANRWLA 1200

QY 1201 VRL5CVGN1CVLPAL1PAV1SRHS1SAG1VGS1SV51LQV1TY1NWLVRMSSEMETN1VA 1260
 Db 1201 VRL5CVGN1CVLPAL1PAV1SRHS1SAG1VGS1SV51LQV1TY1NWLVRMSSEMETN1VA 1260

QY 1261 VER1KEV5ET5EKA1PWO1QETAP5SPQVGRV5FRNYC1RY5ED1D5VLRH1NVT1NGG 1320
 Db 1261 VER1KEV5ET5EKA1PWO1QETAP5SPQVGRV5FRNYC1RY5ED1D5VLRH1NVT1NGG 1320

QY 1321 EKVGI5VGTGAGK5SLT1GLFR1NESAGE511IDGIN1AK1GHLDRFKIT11PODPVL 1380
 Db 1321 EKVGI5VGTGAGK5SLT1GLFR1NESAGE511IDGIN1AK1GHLDRFKIT11PODPVL 1380

QY 1381 SGL1RMN1LDPF5QY5DEEVT51SLAHLKD1F5ALDPK1DHECA5EGEN1SVGQRLVCL 1440
 Db 1381 SGL1RMN1LDPF5QY5DEEVT51SLAHLKD1F5ALDPK1DHECA5EGEN1SVGQRLVCL 1440

QY 1441 ARAL1RRTK1LV1DEA1AAV1D1ETD1L1OST1R1TOPE5DCTV1TAH1N1MTYTR1VL 1500
 Db 1441 ARAL1RRTK1LV1DEA1AAV1D1ETD1L1OST1R1TOPE5DCTV1TAH1N1MTYTR1VL 1500

QY 1501 DKGE1Q5YGA1PSDL1OQ5GLFY5MAKDA5GLV 1531
 Db 1501 DKGE1Q5YGA1PSDL1OQ5GLFY5MAKDA5GLV 1531

RESULT 10
 ABG61810
 ID ABG61810 standard; protein; 1531 AA.
 AC ABG61810;
 DT 15-AUG-2002 (first entry)
 DE Prostate cancer-associated protein #11.
 XX Prostate cancer; prostate tumour tissue; human; mammal; cytosstatic.
 KW Mammalia.
 OS Mammalia.
 PN WO200230268-A2.
 XX 18-APR-2002.
 PD 12-OCT-2001; 2001WO-US032045.
 PF 13-OCT-2000; 2000US-00687576.
 PR 08-DEC-2000; 2000US-00733288.
 PR 08-DEC-2000; 2000US-00733742.
 PR 24-JAN-2001; 2001US-0263957P.
 PR 16-MAR-2001; 2001US-0276791P.
 PR 16-MAR-2001; 2001US-0276888P.
 PR 06-APR-2001; 2001US-0281922P.
 PR 24-APR-2001; 2001US-0286214P.
 PR 30-APR-2001; 2001US-00847046.
 PR 04-MAY-2001; 2001US-0288589P.
 PA (BOSB-) BOS BIOTECHNOLOGY INC.
 XX Glah KC, Mack DH, Wilson KE, Afar D, Hevezi P;
 PI WPI; 2002-471335/50.

DR N-PSDB; ABK92125.
XX Detecting a prostate cancer-associated transcript in a cell in a patient.
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
XX Claim 27; Page 309; 436pp; English.
XX
XX The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridize to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
XX Sequence 1531 AA;
SQ
Query Match 78.5%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSPDLMNNTNTSNDPTKCPONTLVVWPCPYLMACFPFYLXSRH 60
DB 1 MALRGFCSADGSPDLMNNTNTSNDPTKCPONTLVVWPCPYLMACFPFYLXSRH 60
QY 61 DRGYIQMTPLNKTALGFLMIYCMADLFYSFEMERSGIFLAPVFLVSTLIGITTLA 120
DB 61 DRGYIQMTPLNKTALGFLMIYCMADLFYSFEMERSGIFLAPVFLVSTLIGITTLA 120
QY 121 TPLIQERRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180
DB 121 TPLIQERRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180
QY 181 LLLIQVLSCFSRSPLEFSTIHDPNCPRESSASFLSRITFWMTGIVGYKOPLEGSD 240
DB 181 LLLIQVLSCFSRSPLEFSTIHDPNCPRESSASFLSRITFWMTGIVGYKOPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVVKMKKCECAKTRKQPVVYSSKOPAPQPKESKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVVKMKKCECAKTRKQPVVYSSKOPAPQPKESKYDANEVEAL 300
QY 301 IVKSPKEMNPSLFKVLTKTFGYFLMSFFPKA1HDLMSGPQILKLL1KFVNDTKAPD 360
DB 301 IVKSPKEMNPSLFKVLTKTFGYFLMSFFPKA1HDLMSGPQILKLL1KFVNDTKAPD 360
QY 361 WQGYFTVLVFWACQOTLVLAHQYFHCYSGMRKIKAVYGAVYRKALVITNSARKSSTY 420
DB 361 WQGYFTVLVFWACQOTLVLAHQYFHCYSGMRKIKAVYGAVYRKALVITNSARKSSTY 420
QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLMTLGPVSLAGVAVWLVMPVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLMTLGPVSLAGVAVWLVMPVN 480
QY 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYAMELAFKQVLAIRQBELKVLK 540
DB 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYAMELAFKQVLAIRQBELKVLK 540
QY 541 KSAVLSAVGTFVWCTPFIVALCTPAVYVITIDENNTLIDATAVSLAFETILPPLNTLP 600
DB 541 KSAVLSAVGTFVWCTPFIVALCTPAVYVITIDENNTLIDATAVSLAFETILPPLNTLP 600
QY 601 MVLSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGNTSITVNAATFTARSDPT 660
DB 601 MVLSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGNTSITVNAATFTARSDPT 660

DB 601 MVLSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGNTSITVNAATFTARSDPT 660
QY 661 LNTGTFSPRGALVAVVGVGCGKSSLSLTLAEMDVREHVAIKGSVAVPOQAWIQND 720
DB 661 LNTGTFSPRGALVAVVGVGCGKSSLSLTLAEMDVREHVAIKGSVAVPOQAWIQND 720
QY 721 SLRENTILFGCQLEBPYRSYIOACALLPDLEILPSGGRTEIGEGKVLSCGQKORVSLAR 780
DB 721 SLRENTILFGCQLEBPYRSYIOACALLPDLEILPSGGRTEIGEGKVLSCGQKORVSLAR 780
QY 781 AVYSNADIVYFDDPLSADVAHVKHIFENYIGPKGMKNKTRILVTHSMSYLPQVDYIY 840
DB 781 AVYSNADIVYFDDPLSADVAHVKHIFENYIGPKGMKNKTRILVTHSMSYLPQVDYIY 840
QY 841 MSGGKISEMSYQELARDAFAPLRTYASTOEDPAENGTGVSQPKKAKOMENGM 900
DB 841 MSGGKISEMSYQELARDAFAPLRTYASTOEDPAENGTGVSQPKKAKOMENGM 900
QY 901 LVYDSAGKQORQSSSSYSYSDISRHHNSTAELOKAEKKEETWKLMEADKQOTGVKL 960
DB 901 LVYDSAGKQORQSSSSYSYSDISRHHNSTAELOKAEKKEETWKLMEADKQOTGVKL 960
QY 961 SVYMDYMKALGFLISFLSIFLFCNHYVSALASNYWLSLMTDPIVNGTOHTYKRLSVYG 1020
DB 961 SVYMDYMKALGFLISFLSIFLFCNHYVSALASNYWLSLMTDPIVNGTOHTYKRLSVYG 1020
QY 1021 ALGISOGIAVGYSAVSIIGIILASRCLHVDLHSLRSWSPERTPSGNLVNRSKEL 1080
DB 1021 ALGISOGIAVGYSAVSIIGIILASRCLHVDLHSLRSWSPERTPSGNLVNRSKEL 1080
QY 1081 DTVDMSIPEVYIKMFMSLFNVIGACVILATPIAIIIPDLGIYFVQRFVASSROL 1140
DB 1081 DTVDMSIPEVYIKMFMSLFNVIGACVILATPIAIIIPDLGIYFVQRFVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSLDKVDENKAYYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSLDKVDENKAYYPSIVANRWLA 1200
QY 1201 VRLCVCNCIVLFAALPAVISRHSLSAGLVGSYSLOVTTYLNLVRSSEMETNIVA 1260
DB 1201 VRLCVCNCIVLFAALPAVISRHSLSAGLVGSYSLOVTTYLNLVRSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIQETAPSSWPQGRVEFRNYCLRYEDLDPEVLRHNVITNGG 1320
DB 1261 VERLKEYSETEKAPMOIQETAPSSWPQGRVEFRNYCLRYEDLDPEVLRHNVITNGG 1320
QY 1321 EKVGIYGRGAGKSLTLGLFRINESABGEI1IDGINIAKIGLHDLRFKTTIIPDPVLF 1380
DB 1321 EKVGIYGRGAGKSLTLGLFRINESABGEI1IDGINIAKIGLHDLRFKTTIIPDPVLF 1380
QY 1381 SCSLRNMLDPSQYSDSEYVTSLELAHLKDPVSALPDKLDHECAGEGENLSVGOROLVCL 1440
DB 1381 SCSLRNMLDPSQYSDSEYVTSLELAHLKDPVSALPDKLDHECAGEGENLSVGOROLVCL 1440
QY 1441 ARALLAKTKILVDETAADVLETDLLIOSTIRTOEDCTVLTARLNTIMTYRVIYV 1500
DB 1441 ARALLAKTKILVDETAADVLETDLLIOSTIRTOEDCTVLTARLNTIMTYRVIYV 1500
QY 1501 DKGEIOEYGAPSDLQORGLFYSMAKDAQIV 1531
DB 1501 DKGEIOEYGAPSDLQORGLFYSMAKDAQIV 1531
RESULT 11
ID ABM35012 standard; protein; 1531 AA.
XX ABM35012;
AC ABM35012;
DT 08-OCT-2003 (first entry)
XX
XX Cancer based on CYP3A5 related protein seq ID NO:678.
XX

KW Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;
KM cytochrome p450; subfamily IIA; nifedipine oxidase; polypeptide 5;
XX cytosolic.
OS Unidentified.
XX PN W02003013534-A2.
XX PD 20-FEB-2003.
XX PF 23-JUL-2002; 2002WO-EP008219.
XX PR 23-JUL-2001; 2001EP-00117608.
XX PR 24-MAY-2002; 2002EP-00011710.
XX PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX PI Heinrich G, Kerb R;
XX WPI; 2003-268144/26.
XX DR WPI; 2003-268144/26.
XX PT New use of irinotecan for preparation of compositions for treating cancer
XX in subject having genome with variant allele comprising cytochrome p450,
XX subfamily IIA, polypeptide 5 polynucleotide, termed CYP3A5.
XX PT subfamily IIA, polypeptide 5 polynucleotide, termed CYP3A5.
XX PS Disclosure; SEQ ID NO 678; 86pp; English.
XX CC The present invention describes the use of irinotecan (I) or its
XX CC derivative for the preparation of a pharmaceutical composition for
XX CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
XX CC cancer, or malignant glioma in a subject having a genome with a variant
XX CC allele which comprises a cytochrome p450, subfamily IIA (nifedipine
XX CC oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have
XX CC cytosolic activity. The therapeutic applications of (I) is improved,
XX CC since it is possible to individually treat a subject with an appropriate
XX CC dosage and/or an appropriate derivative of (I). Therefore, undesirable,
XX CC harmful or toxic effects are efficiently avoided. Unnecessary and
XX CC potentially harmful treatment of those subjects who do not respond to the
XX CC treatment with substances (nonresponders), as well as the development of
XX CC drug resistances due to suboptimal drug dosing can be avoided. ACFe2200
XX CC to ACFe22751 and ABM34912 to ABM35013 represent sequences used in the
XX CC exemplification of the present invention
XX CC
SQ Sequence 1531 AA;
Query Match 78.5%; Score 7860; DB 6; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCGADGSDPLMDNMTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFYYLYLSRH 60
DB 1 MALRGFCGADGSDPLMDNMTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFYYLYLSRH 60
QY 61 DRGTYONTPLANKTATGALGFLIMVCMADLFYSFMRSGRTFLAPVFLVSPLLGLTTLLA 120
DB 61 DRGTYONTPLANKTATGALGFLIMVCMADLFYSFMRSGRTFLAPVFLVSPLLGLTTLLA 120
QY 121 TELQLERRKGVSSGIMLTFMVLVLCALALIRSKITALKEDAVDLDFDITFYVFS 180
DB 121 TELQLERRKGVSSGIMLTFMVLVLCALALIRSKITALKEDAVDLDFDITFYVFS 180
QY 181 LLLIQLVLSGSDSPLEFSETIHDNPNCPRESSASFLSRTFMWITGLIVRGYROPLESSD 240
DB 181 LLLIQLVLSGSDSPLEFSETIHDNPNCPRESSASFLSRTFMWITGLIVRGYROPLESSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNKKCECAKTRKOPVKKVYSSKDPADPKESSKVDANEVVAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNKKCECAKTRKOPVKKVYSSKDPADPKESSKVDANEVVAL 300
QY 301 IVKSPKEMNBSLRFVLYKTEGPTLMSFFRKATHDLMFSGPOLIKLLIFVNDTKAPD 360
DB 301 IVKSPKEMNBSLRFVLYKTEGPTLMSFFRKATHDLMFSGPOLIKLLIFVNDTKAPD 360

QY 361 WQGYFTYVLLFPDACLQTLVLAHQYFHI CFVSGMRKIKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGYFTYVLLFPDACLQTLVLAHQYFHI CFVSGMRKIKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAORFMDATYINMISAPLOVITLALYLMNLGSPVLAGVAVWVLAHPVN 480
DB 421 GEIVNLSVDAORFMDATYINMISAPLOVITLALYLMNLGSPVLAGVAVWVLAHPVN 480
QY 481 AVNARKTKTYQVAMHMSKDNRIKLMNEILNGIKYLKLYAMELAFKQVLAIRQBELKVLK 540
DB 481 AVNARKTKTYQVAMHMSKDNRIKLMNEILNGIKYLKLYAMELAFKQVLAIRQBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVTIDENNIIADQATFVSLAFNILEPILNP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVTIDENNIIADQATFVSLAFNILEPILNP 600
QY 601 MVTSSIVQASVSLRLRIPLSHELEPDSIERRPVKDGGGNSITVNAFTTARSDPT 660
DB 601 MVTSSIVQASVSLRLRIPLSHELEPDSIERRPVKDGGGNSITVNAFTTARSDPT 660
QY 661 LMGITFSIPGALVAVVGVGCGKSSILSALLAEMDKVEGHVAIKGSVAVYPOQAWIOND 720
DB 661 LMGITFSIPGALVAVVGVGCGKSSILSALLAEMDKVEGHVAIKGSVAVYPOQAWIOND 720
QY 721 SLRENIIIFGCOLLEPYRVSIVQACALLPDLIELIPSGDRTEIGEKVNLGGQKORVSLAR 780
DB 721 SLRENIIIFGCOLLEPYRVSIVQACALLPDLIELIPSGDRTEIGEKVNLGGQKORVSLAR 780
QY 781 AVYNNADIYLPDPLSLVDHVGKHFENYIGPMKKNKTRILVTHSMSTLPQVDVIY 840
DB 781 AVYNNADIYLPDPLSLVDHVGKHFENYIGPMKKNKTRILVTHSMSTLPQVDVIY 840
QY 841 MSGGKISEMSYQELIARDGAFARFKTYASTEBEQDAEENGTVGSGPGKAEAKOMENG 900
DB 841 MSGGKISEMSYQELIARDGAFARFKTYASTEBEQDAEENGTVGSGPGKAEAKOMENG 900
QY 901 LVYDSAGKOLQROJSSSSSYSGDISRRHNSYAEIQAKEAKKETWKLMDKADKAGVYL 960
DB 901 LVYDSAGKOLQROJSSSSSYSGDISRRHNSYAEIQAKEAKKETWKLMDKADKAGVYL 960
QY 961 SVYNDYKAIAGLFISELSIFLPMCNHVSALASNYWLSLWDDPIVNTQOHTKTRLSVYG 1020
DB 961 SVYNDYKAIAGLFISELSIFLPMCNHVSALASNYWLSLWDDPIVNTQOHTKTRLSVYG 1020
QY 1021 ALGISGIIAVFGYMAVSIIGIILASRCLHYDLHSILRSPMSFEERTPSGULVNRFSKEL 1080
DB 1021 ALGISGIIAVFGYMAVSIIGIILASRCLHYDLHSILRSPMSFEERTPSGULVNRFSKEL 1080
QY 1081 DTVDSMIPEVYKMFMSGLFNVIGACIVILLATPIAIIIPDLGLIYFPVQRFYVASSROL 1140
DB 1081 DTVDSMIPEVYKMFMSGLFNVIGACIVILLATPIAIIIPDLGLIYFPVQRFYVASSROL 1140
QY 1141 KRLSESRSRPSYSHFNNTLLGVSVIRAFEEQERFIHOSDLKVDENQAKYPSIVANWMLA 1200
DB 1141 KRLSESRSRPSYSHFNNTLLGVSVIRAFEEQERFIHOSDLKVDENQAKYPSIVANWMLA 1200
QY 1201 VRLCEVNGCIYLPALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRRMSSEMETIVA 1260
DB 1201 VRLCEVNGCIYLPALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRRMSSEMETIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVFRNYCJRYRBDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVFRNYCJRYRBDLPVLRHINVTINGG 1320
QY 1321 EKVGIIVRTGAGKSLTLGLFRINESAGEIINDGINIAKIGLHDAFKTIIIPQDPVLF 1380
DB 1321 EKVGIIVRTGAGKSLTLGLFRINESAGEIINDGINIAKIGLHDAFKTIIIPQDPVLF 1380
QY 1381 SGLSRMLDPPSOYSDDEWMTSLEIAHLKDFVSALPKLDBECAGGEGNLSVGOROLVCL 1440
DB 1381 SGLSRMLDPPSOYSDDEWMTSLEIAHLKDFVSALPKLDBECAGGEGNLSVGOROLVCL 1440
QY 1441 ARALIRTKILLVDEATAVADLETDLLIGSTIRQFEDCTVLTIAHRLNTIMDYTRAVVL 1500

DB 1441 ARAALRKTILVDEAATAVDLETFDDLIQSTIRFQFEDCTVLTIAHLNLTIMDTTRYIVL 1500
QY 1501 DKGEIQEYGA BSDLQQRGLFYSMAXDAGLV 1531
DB 1501 DKGEIQEYGA BSDLQQRGLFYSMAXDAGLV 1531

RESULT 12
ADB20865
ID ADB20865 standard; protein; 1531 AA.
XX ADB20865;
XX
XX
XX 20-NOV-2003 (first entry)
XX
XX MRP1 based cancer related protein SEQ ID NO:678.
XX
XX ifinotecan; colorectal cancer; cervical cancer; gastric cancer;
KM lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KM variant allele; multidrug resistance protein 1; MRP1; cytosolic.
XX
XX Unidentified.
OS
XX MO2003013533-A2.
XX
XX 20-FEB-2003.
XX
XX 23-JUL-2002; 2002WO-EP008200.
XX
XX 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
XX WPI; 2003-354397/33.
XX
XX
XX Use of ifinotecan or its derivative for preparation of a pharmaceutical
PT composition for treating cancer in a subject having a genome with a
PT variant allele comprising a multidrug resistance protein 1
PT polynucleotide.
XX
XX
XX Disclosure; SEQ ID NO 678; 100pp; English.
XX
XX The present invention describes a method for the use of ifinotecan (I) or
CC its derivative for the preparation of a pharmaceutical composition for
CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC cancer, or malignant glioma in a subject having a genome with a variant
CC allele which comprises a multidrug resistance protein 1 (MRP1)
CC polynucleotide (II). (I) has cytostatic activity. (I) or its derivative
CC can be used for the preparation of a pharmaceutical composition for
CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC cancer, or malignant glioma in a subject, where the subject is a human
CC (preferably African or Asian) or a mouse. The present sequence represents
CC a sequence which is used in the exemplification of the present invention.
XX
XX Sequence 1531 AA;
XX

Query Match 78.5%; Score 7860; DB 6; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVNTSNPDFTKCFQNTLVVWVPCFYLMACPFYFLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTVNTSNPDFTKCFQNTLVVWVPCFYLMACPFYFLYSRH 60
QY 61 DRGVIQMTPLNKTALGFLMIVCMADLYSFWERSRGIFLAPVFLVSTLLGITTLA 120
DB 61 DRGVIQMTPLNKTALGFLMIVCMADLYSFWERSRGIFLAPVFLVSTLLGITTLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVALVCALALRSKIMTALKEDAQVLDPRDITFYVYS 180

DB 121 TFLIQLERRKGVSSGIMLTFWLVALVCALALRSKIMTALKEDAQVLDPRDITFYVYS 180
QY 181 LLLIQLVLSGFSRDRSPLESTIHDNCPSSASFLSRITFWMTGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLSGFSRDRSPLESTIHDNCPSSASFLSRITFWMTGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSEOVVPUYKRWKKECAKTRQPKVYVSSKDPQPKSSVVDANEVEAL 300
DB 241 LMSLNKEDTSEOVVPUYKRWKKECAKTRQPKVYVSSKDPQPKSSVVDANEVEAL 300
QY 301 IVKSPKEMNPSTLFKVLYKTFGPFYFLMSFFPKAIHDLMSFGQILKLIFVNDTRAPD 360
DB 301 IVKSPKEMNPSTLFKVLYKTFGPFYFLMSFFPKAIHDLMSFGQILKLIFVNDTRAPD 360
QY 361 WQGYFTVLLFVTAQTLVLHQYFHCIFVSGRIKTAIVGAAYRKALVITNSARKSTV 420
DB 361 WQGYFTVLLFVTAQTLVLHQYFHCIFVSGRIKTAIVGAAYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMISAPLOVILALYLMNLGPSVLGAVAVVWVLPV 480
DB 421 GEIVNLSVDAQRFMDLATYINMISAPLOVILALYLMNLGPSVLGAVAVVWVLPV 480
QY 481 AVNARKTKTYQVAMHMSKDNRIKLANEILNGIKVLYAMELAFKQVLAIROBELKV 540
DB 481 AVNARKTKTYQVAMHMSKDNRIKLANEILNGIKVLYAMELAFKQVLAIROBELKV 540
QY 541 KSAVLSAVGTFTVCTPFLVACTPAVYVITIDENNLDAQTAFAVSLAFNIRPLEP 600
DB 541 KSAVLSAVGTFTVCTPFLVACTPAVYVITIDENNLDAQTAFAVSLAFNIRPLEP 600
QY 601 NVISSIVQASVSLKRLRIFLSHELEBDSIERRPVYDGGGTSITVANAFTMARSPPT 660
DB 601 NVISSIVQASVSLKRLRIFLSHELEBDSIERRPVYDGGGTSITVANAFTMARSPPT 660
QY 661 LINGITFSIPRGALVAVVGVGCGKSSLSALALEMDKVEGHVAKSVAVVPQAWOND 720
DB 661 LINGITFSIPRGALVAVVGVGCGKSSLSALALEMDKVEGHVAKSVAVVPQAWOND 720
QY 721 SLRENILFGQLEBPYRVSITQACALLPDEILPBGDRTEIGKGVNLSCGQKORVSLAR 780
DB 721 SLRENILFGQLEBPYRVSITQACALLPDEILPBGDRTEIGKGVNLSCGQKORVSLAR 780
QY 781 AVYSNADITYLPDPLSADVADHVKHIFENVYIGPKMLKNKTRILVTHSMYSYLPQVDY 840
DB 781 AVYSNADITYLPDPLSADVADHVKHIFENVYIGPKMLKNKTRILVTHSMYSYLPQVDY 840
QY 841 MSGGKISEMGSYOELLARDGAFAPFLRTYASTROBDAENGVTGVSGPKEAKOMENGM 900
DB 841 MSGGKISEMGSYOELLARDGAFAPFLRTYASTROBDAENGVTGVSGPKEAKOMENGM 900
QY 901 LVYDSAGKOLOROLSSSSSYSGDISRHHNSTALQAEAKKETWKLMEADKQOTQVKL 960
DB 901 LVYDSAGKOLOROLSSSSSYSGDISRHHNSTALQAEAKKETWKLMEADKQOTQVKL 960
QY 961 SVYWDYKAIGLFISFLSIFLFMCNHSVASALASYMLSLWTDPIVNGTOGHTYVRLSVYG 1020
DB 961 SVYWDYKAIGLFISFLSIFLFMCNHSVASALASYMLSLWTDPIVNGTOGHTYVRLSVYG 1020
QY 1021 ARGISQIIVFGVSMVSTIGGILASRCLAVDLDHLSLRGMSFFERTPSGNLVNRSKEL 1080
DB 1021 ARGISQIIVFGVSMVSTIGGILASRCLAVDLDHLSLRGMSFFERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPEVIKAFMWSLFFNVIGACIVILLATPAAIIIPPLGLIYFFVQRFVYASSROL 1140
DB 1081 DTVDSMIPEVIKAFMWSLFFNVIGACIVILLATPAAIIIPPLGLIYFFVQRFVYASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEOERFIHQSDLKVDENKAYYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEOERFIHQSDLKVDENKAYYPSIVANRWLA 1200
QY 1201 VRLGCVNCIVCPAALFAVYSRHSLSAGLVGVSVSLOYVTTYLVNMLVMSSMEETNIYA 1260

Db 1201 VRLECVGNCTVLPALFAVVISRHSLSAGVGLSVSYSLQTTVTYINMLVRNSESKEETNIVA 1260
QY 1261 VERLKEVSETEKEAPWQIOETAPSPSPVOYGRVFEFVNYCLAYRBDLPFLVLRHINVTINGG 1320
Db 1261 VERLKEVSETEKEAPWQIOETAPSPSPVOYGRVFEFVNYCLAYRBDLPFLVLRHINVTINGG 1320
QY 1321 EKVGIVERTGAGKSSLTGLFRINESABGEIITDGINIAKTIQLHDLRPKTIITIPQDPVLF 1380
Db 1321 EKVGIVERTGAGKSSLTGLFRINESABGEIITDGINIAKTIQLHDLRPKTIITIPQDPVLF 1380
QY 1381 SGLSRNMLDPPSOYSDSEBVMVTSLELAHKDPVSAIAPKLOHECAGEGENISVGOQOLVCL 1440
Db 1381 SGLSRNMLDPPSOYSDSEBVMVTSLELAHKDPVSAIAPKLOHECAGEGENISVGOQOLVCL 1440
QY 1441 ARALLRKTKILVLEDEATAVVDLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
Db 1441 ARALLRKTKILVLEDEATAVVDLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
QY 1501 DKGEIQEYGAPSDLIQORGLFYSMADAGLV 1531
Db 1501 DKGEIQEYGAPSDLIQORGLFYSMADAGLV 1531

RESULT 13
ADB87954
ID ADB87954 standard; protein; 1531 AA.
XX ADB87954;
XX 04-DEC-2003 (first entry)
XX Human UGT1A1 protein sequence SEQ ID NO:678.
DE
XX irinotecan; cancer; UGT1A1; cytostatic; topoisomerase I inhibitor;
KM colorectal cancer; cervical cancer; gastric cancer; lung cancer;
KM ovarian cancer; pancreatic cancer; malignant glioma;
KM uridine diphosphate glycosyltransferase1 member A1.
XX
OS Homo sapiens.
XX
XX W02003013536-A2.
PN 20-FEB-2003.
XX
XX 23-JUL-2002; 2002WO-EP008217.
PF
XX
XX 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
PA
PI Heinrich G, Kerb R,
XX WPI; 2003-28896/28.
XX
XX
XX Use of irinotecan to treat cancer patient by determining if patient has
PT variant alleles of UGT1A1 gene, administering increased/decreased amounts
of irinotecan based on increased/decreased levels of UGT1A1 gene product.
XX
XX
PS Disclosure; SEQ ID NO 678; 107bp; English.
XX
XX The invention relates to the novel use of irinotecan to treat a patient
CC suffering from cancer. This involves determining if the patient has one or
CC or more variant alleles of the UGT1A1 gene, and if the patient has one or
CC more of such variant alleles, irinotecan is administered in an increased
CC or decreased amount in comparison to the amount that is administered
CC without regard to the patient's alleles in the UGT1A1 gene. The invention
CC has cytostatic activity. A composition of the invention acts as a
CC topoisomerase I inhibitor. The method is useful for treating a patient,
CC an animal e.g. mouse or a human, preferably African or Asian, suffering
CC from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
CC pancreatic cancer or malignant glioma. The present sequence is udes in
CC the exemplification of the invention.

XX Sequence 1531 AA;
SQ Query Match 78.5%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTWNTSNDFTKCFQNTVLVWPVPCFYLMACEPFYLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTWNTSNDFTKCFQNTVLVWPVPCFYLMACEPFYLYLSRH 60
QY 61 DRGIQMTPLANKTKTALGFLMTVYCMADLFYSFWERKRGIFLAPVFLVSPILGITTLLA 120
Db 61 DRGIQMTPLANKTKTALGFLMTVYCMADLFYSFWERKRGIFLAPVFLVSPILGITTLLA 120
QY 121 TFLIQERRKGVSSGIMTFWLVALVCALALIRSKTMTALKEDAOVDLFRDITFYFYFS 180
Db 121 TFLIQERRKGVSSGIMTFWLVALVCALALIRSKTMTALKEDAOVDLFRDITFYFYFS 180
QY 181 LLLIQVLVSCFSDRSPILFSETIHDPNPCPESSASFLSRITFWWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCFSDRSPILFSETIHDPNPCPESSASFLSRITFWWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLYKWKKKCACTRKQPYKVYSSSDPPAPKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLYKWKKKCACTRKQPYKVYSSSDPPAPKSSKYDANEVEAL 300
QY 301 IVKSPQKMNPSLFKVLKTFEGPYFLMSFPFKAIHDLMPFSGPQILKLILKFVNDTRAPD 360
Db 301 IVKSPQKMNPSLFKVLKTFEGPYFLMSFPFKAIHDLMPFSGPQILKLILKFVNDTRAPD 360
QY 361 WQGYFTVLLFVTACIQTIVLHQYFHCIPVSGRIKTAIVAGVYKRALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTACIQTIVLHQYFHCIPVSGRIKTAIVAGVYKRALVITNSARKSTV 420
QY 421 GEIYNLMSVDAORFMDIATTINMTWSAPLOYIILALYLMNLGSPVLAGVAMVLMV 480
Db 421 GEIYNLMSVDAORFMDIATTINMTWSAPLOYIILALYLMNLGSPVLAGVAMVLMV 480
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMEILNGIKYLYAMELAPKDVYLAIRBELVYK 540
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMEILNGIKYLYAMELAPKDVYLAIRBELVYK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTIDENNIIDAOCTAFVSLAFNLRPELNL 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTIDENNIIDAOCTAFVSLAFNLRPELNL 600
QY 601 MVTSSIVQASVSLKRLRIPLSHELEBPDSTERRPVKGGGTNGITVENAFTWARSDDPT 660
Db 601 MVTSSIVQASVSLKRLRIPLSHELEBPDSTERRPVKGGGTNGITVENAFTWARSDDPT 660
QY 661 LINGITFIPRGALVAVVGOVCGGSSLSLALLEMDEVBEHVAIKGSVAIVPOQAWI 720
Db 661 LINGITFIPRGALVAVVGOVCGGSSLSLALLEMDEVBEHVAIKGSVAIVPOQAWI 720
QY 721 SLRENIIFGQLEBPYRSVYQACALPDLIELPSGDRTEIGEGKVMISGGOKORVSLAR 780
Db 721 SLRENIIFGQLEBPYRSVYQACALPDLIELPSGDRTEIGEGKVMISGGOKORVSLAR 780
QY 781 AVYSNADIYLPDDPLSAVDHVAHGKHIFENVYIGPKMLKNTRILIVTHSMGYLP 840
Db 781 AVYSNADIYLPDDPLSAVDHVAHGKHIFENVYIGPKMLKNTRILIVTHSMGYLP 840
QY 841 MSGGKISEMSYQELARDAFAPEFTVYSTQEOEADAEENGVTGVS GPGKEAKOMNGM 900
Db 841 MSGGKISEMSYQELARDAFAPEFTVYSTQEOEADAEENGVTGVS GPGKEAKOMNGM 900
QY 901 LVYDSAGKQOROLSSSSSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKATQGVKL 960
Db 901 LVYDSAGKQOROLSSSSSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKATQGVKL 960
QY 961 SVYWDYKAIAGLFSIFSLFPCNHNVALASNIWLSLWTDPIVNGTQEHYKRLSVYG 1020
Db 961 SVYWDYKAIAGLFSIFSLFPCNHNVALASNIWLSLWTDPIVNGTQEHYKRLSVYG 1020

Db 961 SVYWDYKAIAGLFIISFLIFMCMNVASALASNYWLSMTDDEPIVNGQHTKVLRSVYG 1020
QY 1021 ALGISGIAVFGVSMASISGIIASRCLAHLDHSIRSMSPFERPSSGNLVNRSEKEL 1080
Db 1021 ALGISGIAVFGVSMASISGIIASRCLAHLDHSIRSMSPFERPSSGNLVNRSEKEL 1080
QY 1081 DTVDSMIPVYIKMFMSLFNVIGACIYILATPAAIIIPPLGLIYFVORFYVASSROL 1140
Db 1081 DTVDSMIPVYIKMFMSLFNVIGACIYILATPAAIIIPPLGLIYFVORFYVASSROL 1140
QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQAYYPSIIVANRWLA 1200
Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQAYYPSIIVANRWLA 1200
QY 1201 VRLCEGNCIVLPAALFAVISRHSLSNGVLGVSYSLOTTYLNLVRRSMEMETIYA 1260
Db 1201 VRLCEGNCIVLPAALFAVISRHSLSNGVLGVSYSLOTTYLNLVRRSMEMETIYA 1260
QY 1261 VERLKEVSETEKAPMOIOETAPSSWPQGVAFVFRVYCPRYREDDLPVLRHINVTINGG 1320
Db 1261 VERLKEVSETEKAPMOIOETAPSSWPQGVAFVFRVYCPRYREDDLPVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKIIIPDPVLF 1380
Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKIIIPDPVLF 1380
QY 1381 SSGLRNMLDPFSGYSDEEWTSLFLAHLKDFVSALPKLDHECAEGGENLSVQORQLVCL 1440
Db 1381 SSGLRNMLDPFSGYSDEEWTSLFLAHLKDFVSALPKLDHECAEGGENLSVQORQLVCL 1440
QY 1441 ARRLKRTKILVDEBATAVDLETDLIQSTIRTOPEDCVTLTAHRLNTIMYTRYIVL 1500
Db 1441 ARRLKRTKILVDEBATAVDLETDLIQSTIRTOPEDCVTLTAHRLNTIMYTRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 14
ID ADB96937 standard; protein; 1531 AA.
XX ADB96937;
AC ADB96937;
DT 04-DEC-2003 (first entry)
DE Human MDR1 related protein sequence SEQ ID NO:678.
XX
KM irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KM lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KM multidrug resistance 1; MDR1; cytostatic; human; CYP3A5; MRP1; MDR1;
KM TOPI.
XX
OS Homo sapiens.
XX
PN WO2003013537-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002WO-EP008218.
XX
PR 23-JUL-2001; 2001EP-00117608.
XX
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-268145/26.
PT New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising

PT multidrug resistance 1 polynucleotide.
XX
XX Disclosure; SEQ ID NO 678; 130bp; English.
XX
XX The invention relates to the novel use of irinotecan or its derivative
CC for the preparation of pharmaceutical compositions for treating
CC colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or
CC malignant glioma in a subject having a genome with a variant allele which
CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition
CC of the invention has cytostatic activity. The invention is useful for the
CC preparation of pharmaceutical compositions for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject (preferably human, more preferably African or Asian)
CC or a mouse. The present sequence is used in the exemplification of the
CC invention.
XX
XX Sequence 1531 AA;
SQ
Query Match 78.5%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No.0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSDGSDPLMDMNTVNTSNPDFTKCFQNTVLVWPCFYLMACEPFYFLYSRH 60
Db 1 MALRGFCSDGSDPLMDMNTVNTSNPDFTKCFQNTVLVWPCFYLMACEPFYFLYSRH 60
QY 61 DRGYIOMTPINKTKYALGFLMIVCWADLFYSFWEBSRGIFLAPVFLVSFTLLIGITTLA 120
Db 61 DRGYIOMTPINKTKYALGFLMIVCWADLFYSFWEBSRGIFLAPVFLVSFTLLIGITTLA 120
QY 121 TPLIOERRKGVSSGIMTLFWLVAVCALAIRSKIMTALKEADAQVDPDROITFYVYS 180
Db 121 TPLIOERRKGVSSGIMTLFWLVAVCALAIRSKIMTALKEADAQVDPDROITFYVYS 180
QY 181 LLLIQVLSCFSDRSLFSETTHDPNCPRESSASFSLTFMWTGLIYRGYQPLEGSD 240
Db 181 LLLIQVLSCFSDRSLFSETTHDPNCPRESSASFSLTFMWTGLIYRGYQPLEGSD 240
QY 181 LLLIQVLSCFSDRSLFSETTHDPNCPRESSASFSLTFMWTGLIYRGYQPLEGSD 240
Db 181 LLLIQVLSCFSDRSLFSETTHDPNCPRESSASFSLTFMWTGLIYRGYQPLEGSD 240
QY 241 LMSLNKEDTSEQVPLVKNMKKECAKTRKQPVKVYSSKDPAPKRESSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVPLVKNMKKECAKTRKQPVKVYSSKDPAPKRESSKVDANEVEAL 300
QY 301 IVKSPQKWNPSLFKLYKTRGPFILMSFPFKAHDLMMFSGQIILKLIKPNDRKAPD 360
Db 301 IVKSPQKWNPSLFKLYKTRGPFILMSFPFKAHDLMMFSGQIILKLIKPNDRKAPD 360
QY 361 WQGYFTVLLFVYACQOTVLHQYFHCFSGMRITAVYGAAYRKALVITNSARKSSTV 420
Db 361 WQGYFTVLLFVYACQOTVLHQYFHCFSGMRITAVYGAAYRKALVITNSARKSSTV 420
QY 421 GEIVNLSYDAQGFMDLATYINNIWISAPLOVILALYLLMLNGPSYIAGVAVVWLVPVN 480
Db 421 GEIVNLSYDAQGFMDLATYINNIWISAPLOVILALYLLMLNGPSYIAGVAVVWLVPVN 480
QY 481 AVAMAKTKTYOVAHMSKDNRIKLMNLIINGIKVLYLAMELAFKQVLAIRQEBKLVK 540
Db 481 AVAMAKTKTYOVAHMSKDNRIKLMNLIINGIKVLYLAMELAFKQVLAIRQEBKLVK 540
QY 541 KSAVLSAVGFTFWCTCPFLVALCTPAVYVYTIIDENILIDQOTAFVSLAFNIIREFPLNIIP 600
Db 541 KSAVLSAVGFTFWCTCPFLVALCTPAVYVYTIIDENILIDQOTAFVSLAFNIIREFPLNIIP 600
QY 601 WYISSIVQASVSLKRLRIFLSHEELEPDSIERRPVQGGGJNSITVRNATFTWARSDDPT 660
Db 601 WYISSIVQASVSLKRLRIFLSHEELEPDSIERRPVQGGGJNSITVRNATFTWARSDDPT 660
QY 661 LAGITFSIPEGALVAVVGVGGKSLLSALLAEMKVGHVAKSVAVVPQAVIQND 720
Db 661 LAGITFSIPEGALVAVVGVGGKSLLSALLAEMKVGHVAKSVAVVPQAVIQND 720
QY 721 SLRENTLPGCOLPEEPYRVSIVQACALLPDLLEIPSGDRTEIGKGVNLGGQKQKRVSLAR 780
Db 721 SLRENTLPGCOLPEEPYRVSIVQACALLPDLLEIPSGDRTEIGKGVNLGGQKQKRVSLAR 780

QY 781 AVYSNMDIYLFDDPLSAVDHVGKHIENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYSNMDIYLFDDPLSAVDHVGKHIENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGGKISEMSYQELLARDGAFAPFLTYASTEOBDAENGVTVSGPKAKOMENGM 900
DB 841 MSGGKISEMSYQELLARDGAFAPFLTYASTEOBDAENGVTVSGPKAKOMENGM 900
QY 901 LMTDSAGKOLOROSSSSSYSGDISRHHNSTAELOKA EAKKEEWMKLEADKATQGVKL 960
DB 901 LMTDSAGKOLOROSSSSSYSGDISRHHNSTAELOKA EAKKEEWMKLEADKATQGVKL 960
QY 961 SVYWDYKAIQGLFISFISIFLMCNHVSALASNYMLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYKAIQGLFISFISIFLMCNHVSALASNYMLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQGIANFGSMANVSIGIILASRCLHVDLHSLRSPMSFEERTPSGULVNRFSKEL 1080
DB 1021 ALGISQGIANFGSMANVSIGIILASRCLHVDLHSLRSPMSFEERTPSGULVNRFSKEL 1080
QY 1081 DTVDMSIPEVIKMGMSLFNVIGACIVILATPFAAIIIPPLGIYFFVORFYASSROL 1140
DB 1081 DTVDMSIPEVIKMGMSLFNVIGACIVILATPFAAIIIPPLGIYFFVORFYASSROL 1140
QY 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEBOERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEBOERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVCNCTVLPALFAVISRHSLSAGLVGLSVYSIQVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVCNCTVLPALFAVISRHSLSAGLVGLSVYSIQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEVSETEKEAPMOIOETAPPSMPQVRVEFNVCARREDLPFLRHINVTINGG 1320
DB 1261 VERLKEVSETEKEAPMOIOETAPPSMPQVRVEFNVCARREDLPFLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SSGIRKMLDPPSOYSDBEWTSLBLAHKDFVSLPDKLDHECABEGENLSVGOROLVCL 1440
DB 1381 SSGIRKMLDPPSOYSDBEWTSLBLAHKDFVSLPDKLDHECABEGENLSVGOROLVCL 1440
QY 1441 ABALIRKTKILVLEATRAAVDLETTDILQSTIRIQOFEDCTVLTIAHRLNTIMDTRYIVL 1500
DB 1441 ABALIRKTKILVLEATRAAVDLETTDILQSTIRIQOFEDCTVLTIAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
RESULT 15
ADB92128
ID ADB92128 standard; protein, 1531 AA.
AC ADB92128;
XX
DT 04-DEC-2003 (first entry)
DE Human MDRI related protein sequence SEQ ID NO:678.
XX
KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW multidrug resistance 1; MDRI; cyclostatic; human; UGT1A1; MRP1; Top1.
XX
OS Homo sapiens.
XX
PN W02003013535-A2.
XX
PD 20-FEB-2003.

XX 23-JUL-2002; 2002WO-EP008220.
PF
XX 23-JUL-2001; 2001EP-00117608.
PR
XX 24-MAY-2001; 2002EP-00011710.
PR
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
PA
XX Heinrich G, Kerb R;
PI
XX WPI; 2003-342400/32.
DR
XX
XX
PT New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.
XX
XX
PS Disclosure; SEQ ID NO 678; 104pp; English.
XX
XX The invention relates to a novel use of irinotecan or its derivative for
CC the preparation of a pharmaceutical composition for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject having a genome with a variant allele which comprises
CC a multidrug resistance 1 (MDR1) polynucleotide. A composition of the
CC invention has cytostatic activity. The present sequence is used in the
CC exemplification of the invention.
XX
XX Sequence 1531 AA;
SQ
Query Match 78.5%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTVNTSNPDFTKCFQNTLVVWPCFYLMACFPFYFLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTVNTSNPDFTKCFQNTLVVWPCFYLMACFPFYFLYSRH 60
QY 61 DRGYIQTPLNKTKTALGFLIMIVCAVDLFYSWERSRGIFLAPVFLVSTLLGITTLA 120
DB 61 DRGYIQTPLNKTKTALGFLIMIVCAVDLFYSWERSRGIFLAPVFLVSTLLGITTLA 120
QY 121 TPLIQERRRGVSSGIMLTFWLVALCALAIRSKIMTALKDAQVDLFRDITFYVFS 180
DB 121 TPLIQERRRGVSSGIMLTFWLVALCALAIRSKIMTALKDAQVDLFRDITFYVFS 180
QY 181 LLLIQVLSCFSDRSPLESETIHDNPPCPSSASFLSRIIFWMTITGLIVRGYQPLEGSD 240
DB 181 LLLIQVLSCFSDRSPLESETIHDNPPCPSSASFLSRIIFWMTITGLIVRGYQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKCECAKTRKQPVKYVSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKCECAKTRKQPVKYVSSKDPAPKSSKVDANEVEAL 300
QY 301 IVSPPOKEMNPPLFKVLKTFPGPYFLMSFFPKAIDHLMSSGQIILKLLKFNVDITAPD 360
DB 301 IVSPPOKEMNPPLFKVLKTFPGPYFLMSFFPKAIDHLMSSGQIILKLLKFNVDITAPD 360
QY 361 WQGFYTVLLFWACLOTVLVHOYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420
DB 361 WQGFYTVLLFWACLOTVLVHOYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQREMDLATYINMWSAPLOVILLALYLLMLNLGSLVAGVAVVLMVPVN 480
DB 421 GEIVNLSVDAQREMDLATYINMWSAPLOVILLALYLLMLNLGSLVAGVAVVLMVPVN 480
QY 481 AVNAAKTKTYQVAAHMSKDNRIKLMNIIINGIKYLKLYANIELAKOKYALIRQELAVLK 540
DB 481 AVNAAKTKTYQVAAHMSKDNRIKLMNIIINGIKYLKLYANIELAKOKYALIRQELAVLK 540
QY 541 KSAYLAVGFTFWCTPEFLVALCTPAYVYTIIDENNIIIDAQTAFFSLAFNLIRPPLNLP 600
DB 541 KSAYLAVGFTFWCTPEFLVALCTPAYVYTIIDENNIIIDAQTAFFSLAFNLIRPPLNLP 600
QY 601 MVISSIVOASVSLKRLRIPLSHELBDSIERRPVKDGGTNSITVBNATFTWASDPT 660

Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGJNSITVRNATFTWARSDEPT 660

QY 661 LINGITFSIPGALVAVVGVGCGKSSLSLALLAEMDKVEGHVATKGSVAAYVPOQAMIOND 720

Db 661 LINGITFSIPGALVAVVGVGCGKSSLSLALLAEMDKVEGHVATKGSVAAYVPOQAMIOND 720

QY 721 SLRENILFGCOLLEPPYRSVIQACALLPDLEILPSGDRTEIGEGKVNLSGQOKORVSLAR 780

Db 721 SLRENILFGCOLLEPPYRSVIQACALLPDLEILPSGDRTEIGEGKVNLSGQOKORVSLAR 780

QY 781 AVYNSADIYLFDDPLSAVDHAVGHIFENVIGPKMTLKNKTRILVTHSMSYLPQVDVYIV 840

Db 781 AVYNSADIYLFDDPLSAVDHAVGHIFENVIGPKMTLKNKTRILVTHSMSYLPQVDVYIV 840

QY 841 MSGGKISEMGSYOELLARDGAFAPFLRTYASTEOBDAENGVTGVSQPKAKOMENGM 900

Db 841 MSGGKISEMGSYOELLARDGAFAPFLRTYASTEOBDAENGVTGVSQPKAKOMENGM 900

QY 901 LVTDGAKQLOROLSSSSSYSGDISRHHNSTAELOKAKKEETMKLMEADKAOTGOVKL 960

Db 901 LVTDGAKQLOROLSSSSSYSGDISRHHNSTAELOKAKKEETMKLMEADKAOTGOVKL 960

QY 961 SVYWDYMKAIGLFISFLIFLFCNHSALASNYWLSLMTDPIVNGTOEHTKVRLSYG 1020

Db 961 SVYWDYMKAIGLFISFLIFLFCNHSALASNYWLSLMTDPIVNGTOEHTKVRLSYG 1020

QY 1021 ALGISOGIYAVFGYMANVISIGIILASRCLHVDLHSLRSPMSFFERTPSGNLVNRFSEKL 1080

Db 1021 ALGISOGIYAVFGYMANVISIGIILASRCLHVDLHSLRSPMSFFERTPSGNLVNRFSEKL 1080

QY 1081 DTVDSMIPEYIKMFGSLFNVIACIVILATPIAIIIPPLGLIFFVORFYASSROL 1140

Db 1081 DTVDSMIPEYIKMFGSLFNVIACIVILATPIAIIIPPLGLIFFVORFYASSROL 1140

QY 1141 KRLESVRSPPYSHFNFTLLGVSVYIRAFEBQERFIHQSDLKVDENQKAYPSIVANRWLA 1200

Db 1141 KRLESVRSPPYSHFNFTLLGVSVYIRAFEBQERFIHQSDLKVDENQKAYPSIVANRWLA 1200

QY 1201 VRLECVNCIYLPALPAVVISRHSLSAGLVGLSVYSILOVTTYLNMLVRMSSEMETNIVA 1260

Db 1201 VRLECVNCIYLPALPAVVISRHSLSAGLVGLSVYSILOVTTYLNMLVRMSSEMETNIVA 1260

QY 1261 VERLKEVSETEKEAPWQIOETAPPSWPOVGRVFRNYCLRYREDLDPVLRHINVTINGG 1320

Db 1261 VERLKEVSETEKEAPWQIOETAPPSWPOVGRVFRNYCLRYREDLDPVLRHINVTINGG 1320

QY 1321 EKVGI VGRGTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITTIIIPQDPVL 1380

Db 1321 EKVGI VGRGTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITTIIIPQDPVL 1380

QY 1381 SGSLRNMULDPPSOYSDEEVTSLLELAHKDFVSALPKLDHECAEGGENTL SVGQROLVCL 1440

Db 1381 SGSLRNMULDPPSOYSDEEVTSLLELAHKDFVSALPKLDHECAEGGENTL SVGQROLVCL 1440

QY 1441 ARALLRKTILVDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500

Db 1441 ARALLRKTILVDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500

QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531

Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531

Search completed: December 15, 2005, 15:17:29
Job time : 185.197 secs

This Page Blank (upto)

This Page Blank (upto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:57:04 ; Search time 38.2211 Seconds
(without alignments)
4901.320 Million cell updates/sec

Title: US-10-665-283-8
Perfect score: 10016
Sequence: 1 MALRGFCSADGSDPLMDMNV.....RSVAVAKAKPKFSPDLS 1947

Scoring table: BIOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR 80: *
2: PIR1: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7849	78.4	1531	1 DVHUR	multidrug resistanc
2	4484.5	44.8	1527	2 JE0336	canalicular multis
3	3587	35.8	1545	1 S71841	multidrug resistanc
4	3507.5	35.0	1541	1 S71839	canalicular multidi
5	3289	32.8	1494	2 E89447	protein F57C12.4 l
6	3225	32.2	1573	2 T21219	hypothetical prote
7	3195.5	31.9	1502	2 T42216	multidrug resistanc
8	2838	28.3	1515	1 S51863	cadmium resistance
9	2732	27.3	1398	2 T20434	hypothetical prote
10	2548	25.4	1478	2 T38712	ABC transporter SP
11	2384.5	23.8	1623	2 T01369	ABC transporter At
12	2356.5	23.5	1622	2 D86428	glutathione S-conj
13	2326	23.2	1495	2 E86428	probable ABC trans
14	2312	23.1	1144	2 T27408	hypothetical prote
15	2294.5	22.9	1559	1 S64757	probable membrane
16	2263.5	22.6	1488	2 F86428	probable ABC trans
17	2238	22.3	1516	2 F84919	glutathione-conjug
18	2230	22.3	1539	2 T48059	ABC transporter-11
19	2144.5	21.4	1355	2 T00961	hypothetical prote
20	2144.5	21.4	1514	2 T52080	multi resistance p
21	2140	21.4	1515	2 T52081	MMP-like ABC trans
22	2131.5	21.3	1490	2 T47840	multi resistance p
23	2129	21.3	1545	2 T46645	hypothetical prote
24	2123	21.2	1153	2 T26883	hypothetical prote
25	2109	21.0	1545	2 T42751	sulfonylurea recep
26	2104.5	21.0	1511	2 T42711	sulfonylurea recep
27	2098	20.9	1546	2 T42728	sulfonylurea recep
28	2069.5	20.7	1389	2 T47796	ABC transporter-11
29	2062.5	20.6	1661	2 S64800	probable membrane

30	2037	20.3	1121	2 C87973	protein Y43F8C.12
31	2014.5	20.1	1582	2 A56248	sulfonylurea recep
32	2008	20.0	1592	2 S48933	probable transport
33	1977	19.7	390	2 S68403	inward rectifier p
34	1962	19.6	390	2 J04689	inwardly rectifyin
35	1905	19.0	390	2 A57616	inward rectifier K
36	1902	19.0	390	2 J07901	inwardly-rectifyin
37	1888	18.8	1477	2 S64616	YOR1 protein - yea
38	1884.5	18.8	1037	2 T50518	ABC transporter-11
39	1805	18.0	1146	2 F84487	probable ABC trans
40	1722	17.2	946	1 J05667	multidrug resistanc
41	1693	16.9	1548	1 DVINS	multidrug resistanc
42	1687	16.8	1421	2 T34225	hypothetical prote
43	1606	16.0	1427	2 T20903	hypothetical prote
44	1575.5	15.7	1427	2 T39219	acp-binding caset
45	1558	15.6	1469	2 T50210	probable ABC trans

ALIGNMENTS

RESULT 1

DVHUR
N:Alternate names: multidrug resistance-associate protein (MRP)
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001
C/Accession: A44231; A37495
R/Cole, S.P.C.; Bhargava, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almqvist, K.C.;
Science 258, 1650-1654, 1992
A/Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer
A/Reference number: A44231; MUID:93088080; PMID:1360704
A/Accession: A44231
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: MAPRSGTGMGRGIPATPTSPAPRTSRSSCGIVTSGPV, 50-1531 <CO1>
A/Cross-references: UNIPARC:UPI00001746C; GB:L05628; NID:G1835658
A/Experimental source: small cell lung carcinoma cell line H69AR
A/Note: Sequence extracted from NCBI backbone (NCBIP:119851); this sequence has been co
R/Cole, S.P.C.; Deeley, R.G.
Science 260, 879, 1993
A/Title: Multidrug resistance-associated protein: sequence correction.
A/Reference number: A37495; MUID:93262415; PMID:8098549
A/Accession: A37495
A/Status: not compared with conceptual translation
A/Residues: 1-60 <CO2>
A/Molecule type: mRNA
A/Residues: 1-60 <CO2>
A/Cross-references: UNIPARC:UPI00001746C; GB:L05628; NID:G1835658
A/Note: sequence extracted from NCBI backbone (NCBIP:131929)
C/Genetics:
A/Gene: GDB:MRP
A/Cross-references: GDB:136335; OMIM:158343
A/Map position: 16p13.1-16p13.1
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C/Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmembrane
F/661-844/Domain: ATP-binding cassette homology <ABC1>
F/678-685/Region: nucleotide-binding motif A (P-loop)
F/788-792/Region: nucleotide-binding motif B
F/1310-1503/Domain: ATP-binding cassette homology <ABC2>
F/1327-1333/Region: nucleotide-binding motif A (P-loop)
F/1450-1454/Region: nucleotide-binding motif B

Query Match 78.4%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy	1	MALRGFCSADGSDPLMDMNVNTNTSNDPFRKQNTLVWVPCVLMACPFYLYSRH	60
Db	1	MALRGFCSADGSDPLMDMNVNTNTSNDPFRKQNTLVWVPCVLMACPFYLYSRH	60
Oy	61	DRGYQMTPLNKTATLGLFWIVCWADLFYFWERSRGIFLAPVFLVSPFLGITTLLA	120
Db	61	DRGYQMTPLNKTATLGLFWIVCWADLFYFWERSRGIFLAPVFLVSPFLGITTLLA	120

```
QY 121 TFLIQLERRKGVSSGIMLTFWLVAALCALILRSKIMTALKEDAQVDLFPDITFYVFS 180
D 121 TFLIQLERRKGVSSGIMLTFWLVAALCALILRSKIMTALKEDAQVDLFPDITFYVFS 180
QY 181 LLLIQLVLSCSDSPISFSETHIDNPCESSASFLSITFWMTGLIVRYRQPLESSD 240
D 181 LLLIQLVLSCSDSPISFSETHIDNPCESSASFLSITFWMTGLIVRYRQPLESSD 240
QY 241 LMSLNKEDTSQVVPVLVKNMKCECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
D 241 LMSLNKEDTSQVVPVLVKNMKCECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFRKVLYKTGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPD 360
D 301 IVKSPQKEMNPSLFRKVLYKTGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPD 360
QY 361 MQGFYFVLLFVTAQLOTLVLAHQFHIQFVSGMRITKTAIVGAVYRKALVITNSARKSTV 420
D 361 MQGFYFVLLFVTAQLOTLVLAHQFHIQFVSGMRITKTAIVGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQREMDLATYINMIMSAPLOVILALYLLMLNIGSVLAGVAVMYLMPVN 480
D 421 GEIYNLMSVDAQREMDLATYINMIMSAPLOVILALYLLMLNIGSVLAGVAVMYLMPVN 480
QY 481 AVMAAMKTKTYOVAMKSKDNRIKLMNBIINGIKVILKYAMELAFKDKVLAIRQEBLKYLK 540
D 481 AVMAAMKTKTYOVAMKSKDNRIKLMNBIINGIKVILKYAMELAFKDKVLAIRQEBLKYLK 540
QY 541 KSAVLASVGTFTWCTPPLVALCTPAVAVTIDENNIIIDAQAFSALFNILRPPLNL 600
D 541 KSAVLASVGTFTWCTPPLVALCTPAVAVTIDENNIIIDAQAFSALFNILRPPLNL 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSERRPVKQCGNSTIVRNATFTMARSDPT 660
D 601 MVISSIVQASVSLKRLRIFLSHEBELPDSERRPVKQCGNSTIVRNATFTMARSDPT 660
QY 661 LMGITFSIPREGALVAVVQVCGKSLISALLAEMDKVEGHVAKGSVAVVPOQAMIOND 720
D 661 LMGITFSIPREGALVAVVQVCGKSLISALLAEMDKVEGHVAKGSVAVVPOQAMIOND 720
QY 721 SLRNNILFQCQLEBPYRSVIOACALLPDLLEILPBGDTEIEGKGVNSGGQKRVSLAR 780
D 721 SLRNNILFQCQLEBPYRSVIOACALLPDLLEILPBGDTEIEGKGVNSGGQKRVSLAR 780
QY 781 AVYENADITLFPDDPLSAVDAAVGHKIFENYVIGPKMKNKTRILVTHSMSTLPQVDVITV 840
D 781 AVYENADITLFPDDPLSAVDAAVGHKIFENYVIGPKMKNKTRILVTHSMSTLPQVDVITV 840
QY 841 MSGKISMSGYOELLARDGAPAEFLRTYASTEOBDAENGVTGVSQPGKEAKQEMNGM 900
D 841 MSGKISMSGYOELLARDGAPAEFLRTYASTEOBDAENGVTGVSQPGKEAKQEMNGM 900
QY 901 LVTPSAGQOLOROUSSSSSYSGDISRHHNSTAELOKAAKKEBTYKLMWADKAOGTQYKL 960
D 901 LVTPSAGQOLOROUSSSSSYSGDISRHHNSTAELOKAAKKEBTYKLMWADKAOGTQYKL 960
QY 961 SVYDYMKAIGLFTSFLSIFLPMCHVASALSNWYLSLMTDDPTNGTOEHTKVRLSYVG 1020
D 961 SVYDYMKAIGLFTSFLSIFLPMCHVASALSNWYLSLMTDDPTNGTOEHTKVRLSYVG 1020
QY 1021 ALGISOGIAVFGSNAVASIGGILASRCLHVDLHLSILRSFMSFFERTSPGNLVNFSKEL 1080
D 1021 ALGISOGIAVFGSNAVASIGGILASRCLHVDLHLSILRSFMSFFERTSPGNLVNFSKEL 1080
QY 1081 DTVDSMTPEVITKMGSLFNVI GACTVILLATPIAIIIPGLGLYFFVQRPYVASRQL 1140
D 1081 DTVDSMTPEVITKMGSLFNVI GACTVILLATPIAIIIPGLGLYFFVQRPYVASRQL 1140
QY 1141 KRLSVSSSVYSHHNETLLGVSVTRAFEEQERFTHOSDKVDENOKAYYSIVANRWLA 1200
D 1141 KRLSVSSSVYSHHNETLLGVSVTRAFEEQERFTHOSDKVDENOKAYYSIVANRWLA 1200
```

```
QY 1201 VRLCEVNCIVLPALPAVLSRHSLSAGLYGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
D 1201 VRLCEVNCIVLPALPAVLSRHSLSAGLYGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSMPQYRVEFRNYCYRREDDLFPVLRHINVTINGG 1320
D 1261 VERLKEYSETEKAPMOIOETAPSSMPQYRVEFRNYCYRREDDLFPVLRHINVTINGG 1320
QY 1321 EKVGIVRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFRKTIIPQDPVLF 1380
D 1321 EKVGIVRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFRKTIIPQDPVLF 1380
QY 1381 SGSLRNMULDPSQYSDSEWNTSLELAHAKDPVSAALPKDLHECAGGENLSVQGRQVCL 1440
D 1381 SGSLRNMULDPSQYSDSEWNTSLELAHAKDPVSAALPKDLHECAGGENLSVQGRQVCL 1440
QY 1441 ABALIRKTKILVDEATRAVDLEETDDLIQSTIRIQFEDCVLTIAHRLNTIMDTTRYVL 1500
D 1441 ABALIRKTKILVDEATRAVDLEETDDLIQSTIRIQFEDCVLTIAHRLNTIMDTTRYVL 1500
QY 1501 DKGEIOEYGAPSDILQQRGLFYSMAKXAGLV 1531
D 1501 DKGEIOEYGAPSDILQQRGLFYSMAKXAGLV 1531

RESULT 2
JE0336
canalicular multispecific organic anion transporter - human
C:Species: Homo sapiens (man)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JE0336
R:Juchimul, T.; Hinojilta, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Tsh, S.; Furukawa, M.;
Biochem. Biophys. Res. Commun. 252, 103-110, 1998
A:Title: Isolation of a novel human canalicular multispecific organic anion transporter,
t.
A:Reference number: JE0336, MUID:99032812; PMID:9813153
A:Accession: JE0336
A:Statue: preliminary
A:Molecule type: mRNA
A:Residues: 1-1527 <UCH>
A:Cross-references: UNIPROT:O15438; UNIPARC:UPI0000169859; GB:AF083552
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: ATP
F:11306-1499/Domain: ATP-binding cassette homology <ABC2>

Query Match 44.8%; Score 4484.5; DB 2; Length 1527;
Best Local Similarity 56.6%; Pred. No. 1,56-294;
Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;

QY 8 SAGSDPLMDNVTWNTSNPDFTYCFONTYLVWVPCFYLMACFPFYLYLSRHDRTYQM 67
D 7 SGEIGSKFWDNLSVHTENPDLTFCFQNSILAWPCYILWALPCYLLYLRHHCXYTL 66
QY 68 TPLAKTKALGFLMIYCMADLFYSFWERSGICLAVFLVSPFLILGITTLLATPLIOLE 127
D 67 SHLSKTLKMGVGLVLMCVSWMDLFYFSGLVHGRAPAFVFTPLVVGVTMLLATLLIOYE 126
QY 128 RRGVQSSGIMLTFWLVAALCALILRSKIMTALKEDAQVDLFPDITFYVFSLLIQV 187
D 127 RLQGVQSSGIMLTFWLVAALCALILRSKIMTALKEDAQVDLFPDITFYVFSLLIQV 186
QY 188 LSCFSDSPISFSETHIDNPCESSASFLSITFWMTGLIVRGYROPLESDLSLNKE 247
D 187 LACFREKRPFPSSACNVDPNPIPERISAGFLSLFLFWMTFKMALIYGRHPLEKDWLSLKEE 246
QY 248 DTSQVVPVLVKNMKCECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEALIVKSPQK 307
D 247 DRSQVVPVQQLLEARMKQEKQYARH-----KASAPGK---NASGEDVLLGAPRRP 294
QY 308 EMNLSLFRVLYKTGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPDMQGYRYT 367
D 295 R-KPSFLKALLATGSGSFLISACFKLIQDLISFLNPOLSTILIRPISNMPGSMWGFVLA 353
```


QY	368	VLLPFTACLOTLVYLHOYHICFVSQMBRIKTYAVGVYVRKAVITMSAKRSVGEIVNLM	427
Db	354	GLMFCSMWSQILLDHYTHYIFVTVGKRTGIMGYIRFALVITNSVVRASVGEIVNLM	413
QY	428	SYDARFPMDLATYINMIMSAPLOVLALYLLMLNIGSPSLAGVAVNVLVMPVNAVMANKT	487
Db	414	SYDARFPMDLAPFNLMLNSAPLOIILATYFLMQNIGPSVLGAAMVLLIFLNGAVAKM	473
QY	488	KTYOVAAHMKSKDNRIKLMNEIINGIKVLKLYAMBLAFKDKYLAIRQOEELKVLKKSAYLSA	547
Db	474	RAFQYKMKLKXDSRIKLMSEILINGIKVLKLYAMEPSFLKQVGRIGQEOILRTAVALHT	533
QY	548	VGTFFMWTCTPFLVALCTFAVYVTTIDENNILDAQTFVSLALFNIRFLPNTLIPWYISIV	607
Db	534	TTTFPMWCSPLVYTLITLWVYVYVDPNVNLDAEKAFAVSLSFNIRFLPNTLIPOLISNLT	593
QY	608	QASVLSKRLIRFLSHEELPDSIEBRPVYDGGGTSITVRNATFTMARSDPPTLNGIRFS	667
Db	594	QASVLSKRIQOPLSDEILDPOSVEKTTISPG--YAITIHSTFTMAODLPEPTLHSLIQ	650
QY	668	IPEGALVAVVGVGVGCKSSLALAEADKVEGHVAKGSVAYVYPOQAMIQNDSLRENTIL	727
Db	651	VPKGAIVAVVGVGVGCKSSLVSAALGEMBEKLEGKVMKMSVAYVYPOQAMIQNTCLOENVL	710
QY	728	FGCQLEERYYSVIOACALLPDLLETIPESGDRTEIEKGVNLSGGQKQVSLARAVYSNAD	787
Db	711	FGKALNPKRYOOLTEACALLADLEMLPGGDOTEIEKEGINLSGGQKQVSLARAVYSNAD	770
QY	788	IYLPDPLSADVAHVGKHFENVIGRPKMLKNKTRILVTHSMYLPQVDVITVW6GGKIS	847
Db	771	IFLDDPLSADSHAKHIFDRVIGPEBGVLGKTRVLVTHGISLFPQDFITVLADGOVS	830
QY	848	EMGSYOELLARDGAFAEFLRTYASTEOBODAEENGVTGVSQPKKAKOMENGML-----	901
Db	831	EMGPRPALQNGNSPANFLCNVAPBEDGCHLEDSTALEGADKCALLIEDTLNSHTULT	890
QY	902	---VTDAGKQLOLQOLSSSSSSYSGD-----ISRHNSTAE-LQKAEAKKEETWKLEAD	951
Db	891	DNDPVTYVVOKQFMKOLALS-DEGOGGRPVPRHRLPSEKVOYTFEAKADGA--LTQEB	947
QY	952	KAQNGQVCLSYVMDWKRIIGLFISPLSFLFMCNHSVLSANMYLSLMTDPIYNGTQEH	1011
Db	948	KKAIGTVELSPFMDAKAVGLCTTLATICLLVYGQAAIIGANVWLSAINTNADMSRON	1007
QY	1012	TKVRLSVGALGISQGIAPFGSMANVSIGGILASRCLVNDLHSLIRSPMSFFERTPSGN	1071
Db	1008	TSIRKGVYALGIILOGFLVMLAAMMAAGIOARVLMQALLHAKIRBPQSFDTTPSGR	1066
QY	1072	LVNRSKELDTVDSMIPVYIKFMGSLFNVIACIVILLATPIAIIIPRGLIYFVQY	1131
Db	1068	IINCFSKOIYVYDEVILAVIIMLINSFNAISTLAVINASPLFVVLPIALVITLQY	1127
QY	1132	FYVASSRQKLLESRSRPVYSHNETILGYSVIRAFEOERFIHOSLKDENOKAYP	1191
Db	1128	FYAATSRQKLLESVRSRPIYSHFSEYTGASVIRAYNRSBDFELISPTKADANQSCGP	1187
QY	1192	SIYANRPLAVNLECYGNCIVFPAALFAVYSRHSAGVGVSVSLSLQVTTYLMLNVMS	1251
Db	1188	YIISKRWLSIGVEFPGKCVLFPALFAVYIGSSILNPGVLGVSYSLSLOVYTRALNMIMM	1247
QY	1252	SEMETNIVAVERLKEYSETEKAPWQIOETAPPSMPOVGEVFEFNTCLARREDIDFYLK	1311
Db	1248	SDLESNIVAVERVKEYSKTETEPAPVNVGSRPBGMPRGVEFEPNYSVVRPRGGLDYLK	1307
QY	1312	HINVTINGEKYVIGRTGAGKSSITLGLFRINSEABEGIIIDGINIAKIGLHDLRFKIT	1371
Db	1308	DLSLHVGKEKYGIVGRTGAGSSMTLCLFRLEAKAKEIRIDGLNVADIGLHVRQSQT	1367
QY	1372	IIPDPVYLFSSGLRANLDPFSQYSDPEWYTSLEJLHLDXDFPSALBDXLDHCAEGENYLS	1431
Db	1368	IIPDPVILFSGTILRNNDPFGSYSEBDITWALLESHTLTFVSSQDPAIGDPOCSBGENYLS	1427
QY	1432	VGQRVLCLARALLKTKILVLDEXYAAVDETDLLOSTIRLOFEDCTVITLHARNTI	1491

[illegible]

[illegible]

Db 1467 VLEEDNLLQTTIQLQNEFNACTVITTAHRTIHDSKRWMLDNGKILIEGSPBEELQLPG 1526

Qr 1520 LFYSNAKQAGL 1530

Db 1527 PFYFMAKEAGI 1537

RESULT 4

S71839

canaliculular multidrug resistance protein - rat

C.Species: Rattus norvegicus (Norway rat)

C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

R.Buechler, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, D.

J. Biol. Chem. 271, 15091-15098, 1996

A.Title: cDNA cloning of the hepatocyte canaliculular isoform of the multidrug resistance

A.Reference number: S71839; MUID:96279006; PMID:8662992

A.Accession: S71839

A.Status: preliminary; nucleic acid sequence not shown

A.Molecule type: mRNA

A.Residues: 1-1541 <BUB>

C.Cross-references: UNIPROT:Q63120; UNIPARC:UPI000012578; EMBL:X96393; NID:G1292861; PR

C.Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C.Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein

F.1100-124/Domain: transmembrane #status predicted <TM0>

F.127-151/Domain: transmembrane #status predicted <TM02>

F.160-187/Domain: transmembrane #status predicted <TM03>

F.305-329/Domain: transmembrane #status predicted <TM04>

F.354-381/Domain: transmembrane #status predicted <TM05>

F.431-451/Domain: transmembrane #status predicted <TM06>

F.456-476/Domain: transmembrane #status predicted <TM07>

F.536-564/Domain: transmembrane #status predicted <TM08>

F.574-602/Domain: transmembrane #status predicted <TM09>

F.650-833/Domain: ATP-binding cassette homology <ABC1>

F.667-674/Region: nucleotide-binding motif A (P-loop)

F.966-994/Domain: transmembrane #status predicted <TM10>

F.1018-1046/Domain: transmembrane #status predicted <TM11>

F.1104-1132/Domain: transmembrane #status predicted <TM12>

F.1203-1228/Domain: transmembrane #status predicted <TM13>

F.1313-1506/Domain: ATP-binding cassette homology <ABC2>

F.1330-1337/Region: nucleotide-binding motif A (P-loop)

F.6,1007,1010,1011/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 35.0%; Score 3507.5; DB 1; Length 1541;

Best Local Similarity 46.2%; Pred. No. 26-228;

Matches 721; Conservative 301; Mismatches 471; Indels 69; Gaps 21;

Qr 6 PCSADGSDPLMWNVTWMTSNPDFTKCFQNTVLVWVDFPLMACPF--YFLYLSRHDRG 63

Db 4 FCNS----TFWPLSL--ESPEDLPLCFEQTVLWVPLGLFMLLAPQLVSVSRSTRS 58

Qr 64 YIQMTPLNKTALGLFLWIVCWADLFYSFMERSRGIFLAPVFLVSPFLGLITTLATFL 123

Db 59 SI--TKFYIAKQVFVFLILILAIIDISLALTEDTGQATVPVRVYTNP--TLVLCTWLVLVA 115

Qr 124 IQLERRKGVOSSGIMLT-FWLVALVCAALILRSKIMALKALDEDAVDLFRDITFYVFSLL 182

Db 116 VQHSQWQCMCKKXSWPLSLWILSVLCGVPOQT-LIPLALKDSKSNMAVSYLFFSYSGFQ 174

Qr 163 LIQLVLSGFSRDSPLFSETIHDNCPRESSASPLSRITFMWITGLVGRYQPLSGSLM 242

Db 175 IYLLILITAFSGPS-----DSTQTPSVASFLSITFSWYDRVLKGYGKPLLEBVM 226

Qr 243 SLNK-----EDTSEQVNVPLVKN-----WKKECAKTRKQPVKVVY--SKDPAQKESK 290

Db 227 DIDEQFKTRSVTSKFEAAMTKDLQKARQAFQRLQKSRQRPKATLHGINKKQSOQDVLV 286

Qr 291 VDANEVEEALIVYSPQKKNPSSLFKVLYKTFPGFYPLMSPEFKALIHIDLMFSGPOLIKLLI 350

Db 287 LBEAKKSKKTKTDYIKSM---LIRSLFKTFHVIILKSLFKLIDHLLVFLNPLQLLKLLI 343

Qr 351 KRVNDTKAEWDQGYFTYVLLFVTACLOTIVLHQYFHICFVSGMBRIKTAIVIGAVYRKALVI 410

```

Db      344 GFVSSNSYWFYGCALIMFAVTLIQSFCLQSYFQHCFFVLGMCVRRTVSSYIKKALLTL 403
Qy      411 TNSRKSTVEIYNMVSVDQRFMDLATYINMWSAPLOYTLALMLMAGSVLAV 470
Db      404 SNLRKRYTIGETYNMVSVDQKLMADATNMQVLWSVYIQTLLIFFLMRELGSILAGV 463
Qy      471 AVNVLAVNVAVMAMKTKTYOVAMHKSQNRKIKLMEINLNGIKYLKUYAMELAKDKVLA 530
Db      464 GVMVLILPVNGVLTAKTRINIOVQNMKDKRLKIMNELSGIKLKLPAMEPSPQBOYQG 523
Qy      531 IROBELVLKKSAYLAVGFTWVCTPFLVALCTPAVYVITDENNILLDAQTAFVSLAFN 590
Db      524 IRKRELKMLRFQGLQSLIFILQITPLVSVTFESVYVLVDANVLAERAFSITLFEN 583
Qy      591 ILRPLMLPMTVTSIIQASVSVDRLERYLGDDLDLSAIRVSNPD---KAVFSEAS 639
Db      584 ILRPLMLPMTVTSIIQASVSVDRLERYLGDDLDLSAIRVSNPD---KAVFSEAS 639
Qy      651 FTWASDPPTLNGITFSIPGALVAVVGVGCGKSSLSLALAMDKVEGHVATKGSVAY 710
Db      640 FTWDPDLBATIQDNLDIKPEQLAVAVGVGSGKSSLSVAMLGEMENVGHITTOGSTAY 699
Qy      711 VPQAMTQNDLSRENILFGCOLLEPYRYSVIOACALLPDLLEILPSCDRTGEKGVNLG 770
Db      700 VPQSWTQNGTIKNNILFGSEYNEKXQOVYKACALLPDLLEILFGDMAEIGEGINLSG 759
Qy      771 GQKQVSLARAVSNADYILEFDDPLSAVDHVGKHIFENYIGPKMLKNKTRILVTHSMS 830
Db      760 GQKQVSLARAAVYQDADIYILDDPLSAVDHVGKHIFENYIGPKMLKNKTRILVTHSM 819
Qy      831 YLPQVDTIYVSGCKISEMGSYQELLARDGAFAPLATYAS--TQEDDAENGVTVGSGP 889
Db      820 FLRPVDIIVVAGKTKILEKSYRDLDDKGVFARNMTFTMGSGPEEBATVNN----- 872
Qy      890 GKEAKQENGMVLT-----DSAG-----KOLQROSSSSSYGDISRHNNFAEIO-- 935
Db      873 DSEHEDDDGLIIPMEIIPEDASLARKRENSLRKTSRSRSSSRSGKSLAKSLKIKV 932
Qy      936 ---KAEAKKETWKLMEADRAQOTGOVLSVYWDYKAIQLEIFSLIFLPMCHVSLAS 992
Db      933 NVLKEKEVEGQGLIKKEFETGKVSIFYLKQLAVGMMISLIFILFGLANVAFIGS 992
Qy      993 NYWLSLMT--DDPIYNGI---QEHKVALSVYGAIGISOGIAVFYSNAVSIIGGILASCL 1048
Db      993 NLMWLSATSDSNLNGTNNSSSHRDMRIGVFGALGALOGICLLISTMSIYACRNASKAL 1052
Qy      1049 HVDDLHSILRSPMGFFERTPSGNLVNRFSEKELDTVDMSIPEVIMFGSLFENVIGACIVI 1108
Db      1053 HGQLTNILRAPMRFPTTTPGRIVNRFSGDISIVDDLLPOTLRSMMWCFGIGLIVMI 1112
Qy      1109 LLATPILAIIPPLGLIYFVQRFYVASRQLKLEBSVSRSPVSHFNETLLGVSVIRAF 1168
Db      1113 CMATPVPALIIIPLSIYISVQVFYVATSRQLRDLDSVTSKSPISHSEVYTGPIIRAF 1172
Qy      1169 EEOGRFHQSDLKIDENQKAYPESIVANRMLAVNLECVNGCIYV.PALLPAVISHSISAG 1228
Db      1173 EHOORFLAMNEKQIDINQKCFSWITSNRMFLATLELVGNLVFCSALLIYIYKATLTGD 1232
Qy      1229 LVGLSVSLQVTTTYLMLVMSSEMETNIAVERLKEYSETEKAPMOIQETAPPSWP 1288
Db      1233 VQGVFSLNALNITQTLMLVAMTSSEATNIVAVRISYINVEAEAP--VTDKRPDPMP 1291
Qy      1289 QVQGVREFRNYCLARYREDLDFVLRIHNTVINGEKVGIVRTGACKSSLTGLFPIRESAS 1348
Db      1292 RHGEGIQNNYQVRRPELDVLKGTICNIKSGEKVGVGRIGACKSSLTNCLFRILBSAG 1351
Qy      1349 GEIITDSINIAKGLHLEKRTIIPDDPVLFSGSLAMNLDPSQVDEDEWTSLELAHL 1408
Db      1352 GQIITIDISIVASIGLHLEKRTIIPDDPVLFSGSLAMNLDPSQVDEDEWTSLELAHL 1411
Qy      1409 KDFVSAIPDKLDHCEAGGENTLVGQROLVCLARALRKTKIIVLDATAVADLETDDL 1468

```

```

Db      1412 RSFVSGQLGLSEVTEGQDNLSIGQROLCLGRAVLRSKILVLDATAVADLETDSL 1471
Qy      1469 QSTRROFEDCTVLTIRHRLNTIMDYTRVYVLDKGEIQEAGPSDLLQOGLFYSMAKDA 1528
Db      1472 QTTIRKESQCTVITTAHRLTIIMDSKIMVLDNGKIVEGSPBELLSNNGSFFYLMAKEA 1531
Qy      1529 GL 1530
Db      1532 GI 1533

RESULT 5
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89447
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.el
A:Note: published extrata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: E89447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1494 <STO>
A:Cross-references: UNIPROT:Q20943; UNIPARC:UP1000017801E; GB:chr_X; PIDN:AAA83299.1; P
C:Genetics:
A:Gene: F57C12.4
A:Map position: X
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match      32.8%; Score 3289; DB 2; Length 1494;
Best Local Similarity 47.6%; Pred. No. 1.2e-213;
Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;

Qy      119 LAFILQIERKGVQSSGIMLTFWVALVYCALILRSKIMTALKEDAQVDLFRDITVYV 178
Db      76 LAILTVACNKKGIITSGVITLVWLVVGVGIEFRFYSLGFIYNEVALGIRATLYIA 135
Qy      179 FSLILQIVLSCFSDRSRLPSETIHDPNCPRESSASPLRITFWTGLIVRGYROPLEG 238
Db      136 FTSALHFLFCRAD--VSDMYKSSSCPEYTAGSTINLTQMTGLYILGKSKLEN 192
Qy      239 SDLSINKEDTSQVVPVIVKMKKCECAKTRKQPVKVYSSKDP--AQPRESSKVDANES 296
Db      193 EDLMDNEIDKAKNLIDPSFQNLKPRIDEXH-QNIK-----KQPSALPKN----- 237
Qy      297 VEALIKSPQKKNPSPILFKVLYTTPGPFYLMSPFFRAIHDLMFSGPQIILKLIKPVNDT 356
Db      238 -----HPSFVPIPFYTKYKTLTAGFFYKLCFDMLOFLAPOLIKOLIGFIEDK 284
Qy      357 KAPDMQGYFTVLLPFTACLOTVLHQYPHICVSGRITVAVIGAYRYKALVITNSARK 416
Db      285 NQPMWIGCSIVGIMFSSFLQSMFLHOYHSMRRLGHWASVLTSAVYSRALNINSERK 344
Qy      417 SSTVGEIVNLSVDAQRFMDLATYINMWSAPLOYTLALYLMNLTGSPVLAQVAVVLM 476
Db      345 GKTIQALVNMVSDIQKIDQMAPATIMFWGAPQIPLSIYIFLWKPLGVALAGVVLILA 404
Qy      477 VPVNAVMAAMKTKTYOVAMHKSQNRKIKLMEINLNGIKYLKUYAMELAKDKVLAIROEL 536
Db      405 LPVNGIALIOMRCQTEOMKLDKERIKMSEILNGMKVLKLYSWERSMENMVKIRREL 464
Qy      537 KVLKKSAYLSAVGFTWVCTPFLVALCTPAVYVITD--ENNIIAQAFVSLAFNIRFP 595
Db      465 HIILKLSYFMAALVFSHICAPFLASVYSFVYVYIDPENNVLPETTFVLSFDIIRMP 524
Qy      596 LNLIPVTSIIVQASVSLKRLIFLSHELEPDSIRRRPVKDGCGTNSITVRNATPTWAR 655
Db      525 LAMVAVYGAVQCSVSNTRLKPEFAAEWSPOT---SISHGETSAIIVENGCLFWSMS 580

```



```

QY      668  IPEGALVAVNGVCGCKSLTSLALLAEMDKVEGHVAKGSVAVYVPOQAMIONDSIRENITL 727
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      685  VNRQOLVTVIVRVAAGSSMLQALMGEMKLSGSI SMHGRICVYPODMQWNNLTQRUIT 744
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      728  FGCOLLEBYRVSVAQCALYRDLBELPSGDRTEIGEKVNLGGQKQKVSILARAVSNAD 787
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      745  FGKQFDEYFSRVADACALYRDLQLPLGDNTEIGEKINLSGGQKAKISLARAVYQNH 804
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      788  IYLFDDPLSANDAVGKHIFENVVIGPKMKLKKRILVTHSMSTLPQVDVITVNSGKIS 847
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      805  IYLLDDPMSADVAVHVSQVFGSIVIGPEGMENKTRILVTNLSFLEKSDLIIVNMEGKIE 864
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      848  EMGSYOEILLARDGAFABFL-----RTYASTEQEBODABENGTVGSGGKXAKQEN 898
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      865  YSGKTDLLM--QOGAFEOULLICEKEBERERALEASADEDDENSEBGIMIGDSDFEYD 923
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      899  GMLVTDAGKQLOROLSS--SSYSGDISRHNSTA--ELQKABAKKEETWKLMEA--- 950
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      924  DVM-----ASPIIDHVLGTSHMSTVSGIINRRISTSTHKORRISTTSHSTHSTASTQ 979
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      951  -----DKAQTGVKLSVWDYKALGLFISLISFLMGNHVSALASNYLSLWTD 1003
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      980  TRQLTGERVETGKMDTYKKGAMGMSIAVFLVGMTSTSTFSMGRNMLTDMWSDN 1039
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1004  IV-----NGTOHTKVLISVVALGISOGIAVFGYSMAVSGIILASRLHVDLHSTIRS 1059
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1040  AASGSGTTGQPAIRLGVAVAGLGPSEITLIFIGMLSLTGGVSASRLHAPLMRNLF 1099
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1060  PMSFERTPSGNLVNRFSEKELDYDSMIPVYIKMFCSLPVNIACIVYLLATPIAII 1119
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1100  PMAVYDTTPFRILNRIGKOLETVDLPRVVOGFAQCLQVSTLIIMSTVEGVI 1159
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1120  PRLGLIFPVQRFVVASRQIKLESVSRSPVSHFNETLLGVSVIRAFEBQERFIHQSD 1179
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1160  IPLSVMIWVRXYIATSRLKRLSEITRSPYSHLSIESIGSATIRAYHLVDRFCKLSE 1219
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1180  LKQDENKAYPSPVAVRWLAVLECVGNCVLPALPAVVISRHSLSAGVLGSVSLQ 1239
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1220  TKDSSHQCRKLVNARWMSVRLFEFGNCVLPALPAVTRITTSVGLSVSYALN 1279
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1240  VTYTLNMLVWMSSEMTNIVAVERLKESETEKEAPQIOE--TAPSSMPQGVSEPRNY 1298
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1280  ITTYLNAVAVQITLLEINIVSVERVKYAEETETAEKSEBKPQWNBEGGIWANNY 1339
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1299  CLVRREDLVLRHINVTINGEKVGI VGTGACKSSLTGLFRINSAEGEIIIDGINI 1358
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1340  SARVRPGLNIVVKQLANVEIKRHEKVGIVGRTGAKSSVTLSPRIIAAEGQIVVDGINL 1399
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1359  AKTGLHLRKRITITIPDPVYFGSLRMLNDPESQYDDEWVWLSLELAHLKDFPISALPK 1418
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1400  AELGHLRNLTLTIIPDPVYFGSLRMLNDPFWHYSBGGDITWKLLEMANIKETATANEQ 1459
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1419  LDHECAGGENTSVGROVLVCLARALRKTKIIVLDEATAVADLETDDLOSTRFOED 1478
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1460  LNYITIGSGNISVGOQLVCLARALRKTRVLILDEHTAAVDSITALOKTIREEPAN 1519
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1479  CTYVLTIAHRLNTIMDYTRVIVLDKGEIQEYGAPEDDLQORGL--FYSNAKQAGLV 1531
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1520  ATVLTIAHRLNTIMDYRIIVLNDGKVGFDSPANLLSNRSEFYNAKQAGLI 1573
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1502 <H1R>
A:Cross-references: UNIPROT:O68269; UNIPARC:UP1000012F57A; EMBL:AB010466; NID:93242457,
A:Experimental source: strain Sprague-Dawley; liver
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
Query Match      31.9%; Score 3195.5; DB 2; Length 1502;
Best local similarity 43.1%; Pred. No. 2.3e-207;
Matches 655; Conservative 290; Mismatches 525; Indels 51; Gaps 9,
QY      22  WNTSNPD-----FTKCFQNTVLVWVPCFYLMACFPYFVLVSHRDGYIOMTPIANKKT 75
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      20  MNQTEGPVAVHNLNLCFLAAGSVVPRMVLVGLPYLLXVHNGCCYLRMSLFIKIM 79
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      76  ALGFLMIVCMADLFYSFWERSRGIFLAPVFLVSPITLIGITTLATPLQLERRKQVSS 135
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      80  VLGFLILLYTFNNAVPLMRHGMQPAELLIPVWLMTMSFATFLIHMERKKVGRAS 139
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      136  GIMLTFVLVLYCALIRSKIMTALKEADQVDFRITPVTYSLLIQLVYSCFQDRS 195
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      140  GLLEFGW--LLCCLVPAIDTVQOASGFRQEBLHLATYLLCLVVAELVYSLVDOP 196
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      196  PLFSETIHDNPCESSASFISRTFMWITGLVGRYROPLEGSDLSLUNKEDTSEQV 255
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      197  PFRSEDSKPLNCPPEAASFPKAMFMWAGLKKGRKLGPDKDLSLRENSSEELVS 256
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      256  VLVNWKKECAKTRKQPVKYVSSKDPAPKESKVDANEVALIVKSPQKEMNPSLFX 315
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      257  QLEBERNRN-----FSELPGHGHGSM--GTETETAFL--QERSGQRPBLR 299
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      316  VLKTFEPYPLMSFFFAHDLMMFGSPQLKLKLVNDTKAPDWQGYFTYLLPYTAC 375
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      300  AWRKVFSTPLGLTSLVIDARFAVPKLSLFLBEMGLDESSAMTGMVLAVMFLSAC 359
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      376  LQTLVLYQYHICVSGMRKKTAVIGAVYKALVITNSARKSSVGEIVIMLSVDAQRF 435
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      360  LQTLFEGQYRKVAVLQMRRTAITGLVYKVLVLSGSKRSKSAAGVAVMLVSDVORLV 419
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      436  DLATYIMVSAPLQVILALYLLMLNGPSVLGAVAVMLVAVNVAWANKTKTYQVAHM 495
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      420  ESTIHNGMLFLFWIIVCVYIMQLGPGSALTAVAVFLSLPLNFTITKRSFHQEO 479
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      496  KSKDNRIKLMELINGIKVLYLAWELAFQKVLATIQEELKVLKKSAYISANGTFPWC 555
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      480  ROKASRARLTSSMLRTVRTIKSHGWECAFERLHTRGQELGALKTSAFLSVSVFOV 539
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      556  TPLVVALCTPAVYVTTDENNIILDAQAFVSLALFNILRFPNLTPMWISSIVQASVLKR 615
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      540  STELVAVLVRAVHTLVADNDAMDARFVTLTVLSTLNKAQAPFESVHCLVQARVSFDR 599
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      616  LRIFLSHELEPDSIERRPYKDGSTNSITVRNATFWARSDDPTLNGITFSPREGALVA 675
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      600  LAFLFLCLEBVDPRGMVLSPPR--CSSKDRISIHNGTFMSESPCLHGMVLYPOGCLLA 658
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      676  VVGQVGGKSSLLSALIAENDKYEGHVALIGSVAYVYVPOQAMIONDSIRENITLFGCOLLEP 735
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      659  VVGPFVGGKSSLLSALIGELIKYEGSVSIEGSAVYVPOEAMVONTSVENVCFOEIDL 718
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      736  YRSVSVIOACMLDLDELTPSGDRTEIGEKVNLGGQKQKVSILARAVYNAADVILPDP 795
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      719  WLQEVLEACALGSDVASFPAQVHTPVGEQGMNLSGGQKQKSLARAVYRRAAYLMDPL 778
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      796  SAVDAVGHKAI FENVVIGPKMKLKKRILVTHSMSTLPQVDVITVNSGKISGYSOEL 855
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      779  AALDAVHVSQEVFOVIGPSGLLGTTRILVTHLHVLPQADQLIVLANGTIAEMGSQDL 838
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      856  LAADGAFABFLRYASTEQDAENGVT-----GVSGGKAKQKQENGLVYTDAGKQ 909
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      839  IHNHNGALVGLLDARQPAGEGBEAAHAATSDDLIGSGSGGTFTRPPE----- 886
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      910  LQRLSSSSGYSQDISRHHNSTAEELQKAEKKEETWKLMEADAKQGVGVGLVYWDYMK 969
      :  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


[illegible][illegible]

Db 1088 DSISRPYYAAGFEALNLSITIRAYKADNRADINGRMDNNINRFTLVNMGANRWLGRL 1147

Qy 1204 ECVANGCIVLFAALPAVI-----SRHSLSAGLVGSVSYSLQVTTYLANLVKSSMETN 1258

Db 1148 ETLGGLMIWLFAASFVAMONGRAENQAFASTMTGILLSYALNITISLTITGLRLASLAENSL 1207

Qy 1259 VAVERLKEYSFEKKEAPWQIQIETAPSSMPQVGRFEFNNYCLARRDEDFYLRHINTYN 1318

Db 1208 NAVEKVGNYIIPPAAPVIEINNRPMPGMPSSGSIKFEEDVVLRYRPPQLPVLHGVSEPIH 1267

Qy 1319 GGEKVGIVGRTGAGKSSLTTLGLFRINESAEGEIIIDGINAIKGLHDLRFKITTIIPOPV 1376

Db 1268 PTDKVGIVGRGAGKSSLLNHLFRIVAEKGRITLIDDCVGRFGMLDRKVLGIIPOSPV 1327

Qy 1379 LFGSGSLRNMLDPFSQYSDDEBWTSLLELAHLKFDVPSALPDKLDHCEAGEGENISVGORLV 1438

Db 1328 LFGSGVNRNLDPPFGHNADLMESIERAHLKDTIRRNVLGIDABVSENGENFVSQGRLL 1367

Qy 1439 CLARRLKKILVLDEATAAVDETDDLQSTIRTOGEDCTVLTIAHRLNTMDYKVI 1498

Db 1388 SLSLRALRSKILVDEATAAVDVTDLIQITIREEKSCMTMLIAHRLNTIIDCKIL 1447

Qy 1499 VLDKGEIOEYCAPSDLLQORGLFYS-MAKDAG 1529

Db 1448 VLDSGRVGEFSSPENILSNEGSSFEKMWQSTG 1479

RESULT 12

D864428
glutathione S-conjugate transporting ATPase (AtGMP1) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D864428
R:Thelloglis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chih, C.W.; Chung, M.K.; Corn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:1130712

A:Accession: D864428
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1622 <STO>
A:Cross-references: UNIPROT:Q9CGG9; UNIPARC:UPI000009CF47; GB:AE005172; NID:g11055814; F
C:Genetics:
A:Map position: 1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match	23.5%	Score 2356.5	DB 2	Length 1622
Best Local Similarity	35.6%	Pred. No. 1.8e-150		
Matches	569	Conservative 274	Mismatches 564	Indels 193
				Gaps 31
QY	47	WACEPF-FEYLISRDRGYIQWTPLNKTKTALG-----FLMIVCMADLFYSFM	94	
DB	8	WYCKEVPNGWTKTVDYAFGAVTPCAIDSPVLGSHVLVLICLRMLWITKDHVDFEC	67	
QY	95	ERSR-GIFLA-----PYFLVSPILLGITT-----LATFLIQLEKRVQVS	134	
DB	68	LRSKFESYFLALAAVATAPLFRU--VMRTSVLDLDGAGPPYEAEMLVLE--AFRWG	122	
QY	135	SGIMLTF-----MLVALCALAIISKINKITALKEDAOVLDFFDITYEYVFSLL	183	
DB	123	SALVMTVETKTIHILRMVYRFAVITYALVGDVMLNLVLSVK-EYGSFKYLILISEVA	181	
QY	184	IQLV-----LSCFSDRSPLFSETIHD-----PNCPESASASTLSNITWW	223	
DB	182	VQVAFGTLIPVYFPNLDIPGYTIPVGTJENSEBEYEYELPGGENICPERHANNIFDSIFPSW	241	

QY	224	ITGLIVNGYQPLEGSDMLNEDTSEOVVPLVKNWKECKATRKQOPKXVYSSDPA	283
Db	242	INPLMTJSGKRPLETEKOVMLDWDKTEJLNRSFQKSWDXELEK-----	285
QY	284	QPKESKVDANEEVEALIVKSPCKENMPSLFKVLYKFKGYPILMSPEFKAIHIDIMESGP	343
Db	286	-PK-----PWLLRLNNSLGGRFPMWGKJGNDOSQVGR	320
QY	344	QILKLIKFNVDYKAPDMQGYFYTVLLFVTACIQTIVLHQYFHIQVSGMRIKTAVIGAV	403
Db	321	LLNLBELKSM-QLNBPAMIGIYAIISFVGVLGVLCEAQGFQVWMBVGRILRSALIAAV	379
QY	404	YRKALVITNARKSSTYGEIYVNLMSVDQAQRFMDATYINIMSPLOYIALLYMLMLNG	463
Db	380	FRKSLRLTNEBKRFQTKITLMTTAAESIQQICSLHMTMSAPFRRIYALVALLYOQG	439
QY	464	PSYLAGVAVMWLPVAVNAVMAMKTKTVQVAMKSKDRIKMLNMLNGIKULKYAMELA	523
Db	440	VASITIGALFLVLMPIQTVIISKTQKITEGLQDTDKRIGLMEVLAAMDYKCYANENS	499
QY	524	PKDKYLAIRQBELKVLKKSAYLSAVGTFTWCTPPLVALCTFAVYVTTIDENNILDAQTAF	583
Db	500	FQSKQVTRDDELSPFRKAQQLSAPNFMIINSIPVLTVVSVFGVFSLLGGD--LTPARAF	557
QY	584	VSLALFNIRPLNITLPMVJISSIYQASVSLKRLIRFILSHEE--LEPISERRPKYDGGT	641
Db	558	TSLSLFSFLRPLMLPNIIITQWNAVNSINRLEEVLSTERVLLPNP---PLEP--GQ	611
QY	642	NSITVRNATFW-ARSPDPTLNGITFSIPEDALVAVVGVQCGKSLLSALLAEMD--KVE	699
Db	612	PAISIRNGYFPMDSKADRFILSINLNDIPLGSLVAVVGSYBEGKTSLSIAMLGSLPARSD	671
QY	700	GHAIVAKGSVAVVPOAMIONDSIRENITLFGQOLEEPYRSVYQACALLPDLLEILPSGDR	759
Db	672	ATVTLRGSVAVVPQVSMIPATVADNITLFGAPRQEKXERIVDTALQHDLELLPGDILT	731
QY	760	EIEGKCVNLSSGGQQRVSLARAAYSNAADYILPDPPLSANDAHVAKHIFENVYIGKGMULN	819
Db	732	EIEGRGVNISSGGQQRVSMARAAYVSNBVCILDPRLPALDAHVOQOFEKCI--KREIQQ	789
QY	820	KTRLIVNHSMYLPOVNVITIMSGGKISSEMSYQELLARBDGAFAPLRTVASTEQEDAE	879
Db	790	TTRVAVINQHLFSLQVQDKILVHEGYKEBTTYELCHSGPLFRMLBNACKVEDY--SE	847
QY	880	ENGVYGVSGPKEKAKOMENGLVYDSAQLOROLSSSSSYSGDISRHNNSTAELOK--	936
Db	848	ENGEAEVD--QTSXKXPENG-----NANLQDGI	875
QY	937	-AEAKKEETWKLEADKAOQGOVLKYVWDYMKAI-GLTFSLSIFLPMCHVSALASNY	994
Db	876	ETTKSKEGNSVLVRREREITGVSMKYLERYQNLAGAMVMMVLIVICVLTQVPRVSSST	935
QY	995	WLISLMTD--PIVNGTOEHTKVRLSYVGCALGISGIAVFGVSMABVSGILASCLVYLD	1052
Db	936	WLSEWTSOSGTRKTHGPLFYN---IYVALLSFGQVSVTLINSYVLLMSLSYLAACKMDAM	991
QY	1053	LHSIIRSPMSPEFERPTSGCNLVNRPSEKELDTVDSMIPEVIKFMGSLFVFIACIVILAT	1112
Db	992	LGSIIIRAPMVFOQNPBLGRINNRKAKOMGQIDRIVAVFVNMFKSIAQOLSTVILIGVS	1051
QY	1113	PIAAIIIPGLIYFVQRFYVASSROLKRLSVSRSPVSHFNETLLGVSVIRAFBEOE	1172
Db	1052	TLSLMAIIMPLVIFYGAVLYQNTSREIKRMDSTTRSPVVAQFGEALNGJSSIRAYKAD	1111
QY	1173	RFIHOSDLKVDENOKAYIPSIIVANRWLAVRECCYGNCTVLPALPAYI-----SRHSISA	1227
Db	1112	RMAEINGRSMDDNIRFTLVNNAARWKGIRLEVIYGGIMVWLTAFLAWQNGKAMQOAYA	1171
QY	1228	GLVGVSVSYQVTVTYNMLVMSSEBMTNVAVERLKEYSETSEKAPMOIQETAPSSW	1287
Db	1172	STMGLLISYALSTISSLAVYRLRLSLAENSLANSVERGANTYIETSEAPLVYIENNRPPPGW	1231
QY	1288	POVGRVEFRNYCLYREDBLPVLRIHINVITNGEKVGIQVGTGAKGSLTGLFRINESA	1347

```

      1232 PSSSIFKEDVAVLRRELPEVLHGVSFLISPMDKVGVIGRTGAGKSLNALPRIYELE 1291
      1348 EGEIILIDGINIAKGLHDLRKITIIIPQDPVLFSGSLAMNIDPFSQYSDSEWMTSLIAH 1407
      1292 KGRLLIDECIDIGRGLMDLRKVLGIIPAPVLFSGYVAFNIDPPSEHNDADLWESLERAH 1351
      1408 LKDFVSLPDKLDECAEGENGLSVGQOLVCLARALLRKTIIVLBETAADVLETPDL 1467
      1352 LKQIRRNPLGLDIEVTAGENFSGQRLSLRALIRKSLIYLBETAADVLETPDL 1411
      1468 IQSTIRQFEDCVLTITAHRLNTIMDYTRVILDKGEIOEYGAESDILQGRGLFYSMAKD 1527
      1412 IQKIRREFKSCVTMLIIAHLRNTIIDCKVLVDGKVOEFSFSPENLISNESSFSK--- 1468
      1528 AGLVGGGGGMLSKRGKIIEPVYVLTRLAEPAERIRYTRRE 1567
      1469 -----MVOSTGTANMEYLSITLENK-----RTRE 1493
  
```

RESULT 13

```

      E86428
      Probable ABC transporter [Imported] - Arabidopsis thaliana
      C/Spectes: Arabidopsis thaliana (mouse-ear cress)
      C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
      C/Accession: E86428
      R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
      Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
      ansen, N.F.; Hughes, B.; Hultzer, L.
      Nature 408, 816-820, 2000
      A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
      C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marziani,
      Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
      A/Authors: Salzberg, S.W.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
      Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
      A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
      A/Reference number: A86141; MUID:21016719; PMID:11130712
      A/Accession: E86428
      A/Status: preliminary
      A/Molecule type: DNA
      A/Residues: 1-1495 <STO>
      A/Cross-references: UNIPROT:Q9C8H0; UNIPARC:UPI0000482FF; GB:AE005172; NID:g11055818; F
      C/Genetics:
      A/Map position: 1
      C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
  
```

```

      Query Match      23.2%; Score 2326; DB 2; Length 1495;
      Best Local Similarity 34.3%; Pred. No. 1.8e-148;
      Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;
  
```

```

      QY 2 ALMGFGC--ADGSDPLMDANVTNNTSNPDFTKCPQNTVLVWVP-----CFY-LMAQCP 51
      DB 5 ALMWYCPVADGF-----WEKAVDGAFAAYTPCAIDSLVMSHFVLLGLCFYRIWTF- 58
      QY 52 FYFLVLSRHDGXYIQMTPLKTKTALGFLMTIYQWAD-----LFYSFMRSGIFLAPV 105
      DB 59 -----HNKKAQIYVRKKYINCVLG-LACYCVAEPVRLVWGLSLFMDDETPDPF 110
      QY 106 FLVSPVLLGTTLLATFLIQLERRKQVSGSGLMTFLVALVCALILRSKIMATLKEDA 165
      DB 111 EVASLWVEAFWAFWSMLVLIGLETKQYKEFRWYVRFVLVAVADVLVLPLKNSIN 170
      QY 166 QVULFADIT---FYVYFSLILIQVLVSCFSRDRSPLSETTHDP-----NPPC 209
      DB 171 RYALYFISSRGQALFGILLIYIPEL--DPYFGYIVNNEPLDNEVDALNGGEIICP 228
      QY 210 ESSASLSRITFMWITGLIVRGYRQPLEGSDLSLNEKEDTSEQVAVLVKMKKECAKTR 269
      DB 229 ERASISFSRIYFWMITFLMGLGRPKRTEDWQDLKMDQTEFLIKRFGSCWIBESR-- 286
      QY 270 KQPVKVVYSKSDPAQPKESKVDANEVEALLVSKPKQKPNPSLFKVLYTKFPYPLMSF 329
      DB 287 -----PK-----PWLLRALNNSLGGRFWLAG 307
  
```

```

      QY 330 FFKAHDLAMFSGPQILKLKIPVNDIKADPMQGYFTVLLFVTACLOTVLHGYFNHCF 389
      DB 308 IFKIGNDLSQVGVVILSHLSRQEGED--PAWGYVAFLIFVGVTLGVCEAYFQNWV 366
      QY 390 VSGKRITAVYGAAYRRKALVYTNARSARKSTVGEIVNLSMDAOCFMDLATYNNIWSAPL 449
      DB 367 RVGRRLASTVVAALFHSLSLTHEARQNPASGKVTMTITDANLQOISQHLMSARPF 426
      QY 450 QVIALYLLMLNIGPSYLAQVAVNVLVNVAVNMAKTKTYQVAHMSKDNRIKLMNEIL 509
      DB 427 RIIVSMILLYQQLGVALSFLSLLIFLILPIQTLISHRKLTREGLQMTDKRVGITNEIL 486
      QY 510 NGIKVLYLAMELAFKQKVALIROEELKVLKSAVYLSAVGTFVTCFPLVALCTPAAVY 569
      DB 487 SSMDTVACVAMEKSFESRIGINNEBELSPRQQLSAFNSPLNISPVVATVVSFGV 546
      QY 570 TIDENMLDAQTAFAVSIALFNILRFPNLIIPMVSSIVQASVSLKRL-RIFLSHELEPD 628
      DB 547 LIGGD--LTPARAFTSLFAVLRFPNMLPNLSQVYVNAVSLQRELLISERILAQ 604
      QY 629 SIERRPVKGGGNTSITVRNATFTW--ARSDPPLNGITFSIPGALVAVVQVQCGKSL 687
      DB 605 NPLQOP-----GTPAISIKNGYFSDSKTKPTLSDINLEIPVGTVAIVGTEGKSTL 659
      QY 688 LSLALLMDKVE--GHAIVKSVAVVPOQAMIQNDSLENTLFGQLEPPTRYIOCAL 746
      DB 660 ISMLGLSLHAETTSVVRKSVAVVPOVSVIFNATVRENTLFGSDFESERYRAIDATYAL 719
      QY 747 LPLELILPSGDRTEIGEGKYNLSGGQKORVSLARAVVSNADIVLPDPLFANDVAHVAKHI 806
      DB 720 QHDDILPLGDLTEIGERGYNISGGQKORVSMARAVVNSDVYIFDPLSALDHAHVQV 779
      QY 807 FENVIGKGMKNKTRILVTHSMSTYLPQVDVVIYMSGKISENGYSOELLARQAPAEFL 866
      DB 780 FDSGM--KDBLRGKTRVLVTLNQHLPLMDKIIIVSGMCKEBSGTFELSKSGLFKKLM 837
      QY 867 RTYASTBQEDDAEENGVTGVSFGPKAKQKQENGLYTDGSKQLOLQSSSSYSGDISR 926
      DB 838 ENAGKRDATQEVNTN-----DENILKGLPPTVVDVS-----ERNLGSSTK----- 876
      QY 927 HHNSTABLQAEAKKETWELMEADRAQOTQVLSYWDYMKAI-GLFISFLSIFLPMCN 985
      DB 877 -----QGRKRS-VLIKQERETGIIISNVLAKYKAVGGILVMMILACTYAT 924
      QY 986 HVSALASNYWLSMTDPIVNGTOEHTK-----VRLSVYAGLGISQIAVGYSMVSGIS 1040
      DB 925 EYLRVSSSTWLSIWTD-----QSTSKVYSPGYIVVVALDGFQVAVFTMSFMLITS 977
      QY 1041 GILASRCLHVDLHSLIRSPMSFFERTPSGNLVNRSKELDYDSMIPVYIKRFGSLFN 1100
      DB 978 SLHAARLHDMASSSLRAEMLFPHNPTQVRVNRPSKIDGIDRRVAVANIMNFMNQLMQ 1037
      QY 1101 VIGACTVILLAPRIAIIIPGLIYFPVQRFVVAASROKRLSEVSRSRPSVSHENETL 1160
      DB 1038 LLSFTPLIGVSTISLMAIPLELILFYAAVLYQSTSRERKRLDSVTRSTIYQFGBALN 1097
      QY 1161 GVSIVRAFEROEERFIHQSDLKVDENOKAAYPSIVARMLAVRLDECQNCVLFAALFAVI 1220
      DB 1098 GUSISIRAYAYDMARKNGSMONNIRFTLANSSNRMLIRLETIGGWIMWLATPAVL 1157
      QY 1221 SRHSLS-----AGLVLSVYSIQVTTYLVNMLVRMSSEMETNIVAVBRLKEYSETEKEAP 1275
      DB 1158 QNGNTNNOAGFASTWGLISLTYLNTISLSGLVROASRAENSINSVERVGNVYIDLSEAT 1217
      QY 1276 MQIQETAPSPSQVQVRFVFRNYCLRYREDLDVLAHINVTINGKEVGVVIGRTGAGKSS 1335
      DB 1218 DLIENNRPVCGWSPGSGSIFEDVHLAKYRGLPPLVGLITFPVSPSEKVGAVGTGAGKSS 1277
      QY 1336 LYLGLFRINESAGEIILIDGINIAKGLHDLRKITIIIPQDPVLFSGSLAMNIDPFSQYS 1395
      DB 1278 MLNALPRIIVEGKRIMIDCDVAKFGLTDVRVLSIIPQSPVLFGTVRFNIDPSEHN 1337
  
```

[illegible]

Db 530 ---KSEKPKQKEKAVE--NVQGRKVKSVYQYIKITMGFIENSSAPLILFIAHTVIMNR 584
QY 993 NYWISLWTD-----PIYNGIOE-----HTKVLSTYGAIGISOGIAVFGYSM 1035
Db 585 SLMLSDMSNENAIKKATLSSVDYLNSTSSVDGVSFETRLIYACG---GLEMLLTLAL 641
QY 1036 A---VSIIGLIASRCHVDLHLSILRSPPMSFFERTPSGNLVNFRSKLDYDMSIPEYK 1092
Db 642 AFTVTLTGLSRASYGHSPLIHALLVAPISFFPTTPTGRILNIRSLRDIVDK--LQDNIR 700
QY 1093 MEMSGLFNVIAGACIVILLATPIAAIILPPLGLIYFFQRFYVASSRQDKRLSESRSBPV 1152
Db 701 MCTQTLNACMLIVLISITPIFLVCAAPILILYFMIIYIPIRSQDKRLSANSRPIIL 760
QY 1153 SHFNETLLGVSVIRAEFEGERFIHQSLDKVDENOKAYPSIVANRWLAIRIECVGNCTVL 1212
Db 761 STIASIHGASSIRAFDKTERTTTALSTVNDKFAQCRKYLISHMSNRMLATRLLELGNFCVL 820
QY 1213 PAALFAVIVSR--HSLSAGLVGSVSYSLQVTTYTLNMLVRRSSSEMTNIVANERLKEYSET 1270
Db 821 FASLTATYSTKYEGLTPGAGLSVSYALTITTEVLANICVRVSEIESNINVSERVNEVYKL 880
QY 1271 EKEAPMOIETAP--PSSMPQVGFVEFRNYCLARREDLDFLRIHNTVINGEKYIVGRT 1329
Db 881 EPEARWRIRKSLSENEKMWVYKKTIEDGFSMRIRKNIPLVTKNIDLKEGGERIGVIGRT 940
QY 1330 GAGKSSLTGLGRFINESAGEIILIDGINIAKIGLHDLRFKITIIPDPVLFSSGLRMNLD 1389
Db 941 GSGKSSLTVALYRMITEGESGTIKIDVEIDPTIGLHQRSKLIIIPQPVVPSGTLRFNLD 1000
QY 1390 PPSQVSDDEVWTSLELAHLKDFVSALPDKLDHCAEGEMLSVQORQVCLARALLRRTK 1449
Db 1001 PENQVSDQDQIWNCLCQLKQFAQEDDKTLDRYIAEGKMSVGERQLCLCRALLRKRAR 1060
QY 1450 ILVLEATAVAVLETTDDLLQSTIRTPQEDCTVTIARLNTIMDYTRIYVLDKEIOEYX 1509
Db 1061 IYILBEATIASVPTVTVDIGVQRAIRQHPPOSTTISIAHRLDTIYDSXKIVLVDAGRVAEFD 1120
QY 1510 APSDLLQOGRLEFYS 1523
Db 1121 TPSNLLNPDLSLYS 1134

RESULT 15
S64757
Probleme membrane proteien YIL015w - Yeast (Saccharomyces cerevisiae)
N.Alternate names: hypothetical protein A255; hypothetical protein L1313
C.Species: Saccharomyces cerevisiae
C.Date: 10-Sep-1999 #sequence Revision 10-Sep-1999 #text-change 09-Jul-2004
C.Accession: S64757; S64763; PS0041; S69391; S70560
R.MiOga, T.; Zimmermann, F.K.
submitted to the Protein Sequence Database, May 1996
A.Reference number: S64743
A.Accession: S64757
A.Molecule type: DNA
A.Residues: 1-1559 <MO>
A.Cross-references: UNIPROT:PI4772; UNIPARC:UPI0000126A9C; EMBL:Z73120; NID:g1360184; PI
A.Experimental source: strain S288C
R.Goffeau, A.; Purnelle, B.
submitted to the Protein Sequence Database, May 1996
A.Reference number: S64761
A.Accession: S64763
A.Molecule type: DNA
A.Residues: 1-1559 <GO>
A.Cross-references: UNIPARC:UPI0000126A9C; EMBL:Z73120; NID:g1360184; PIDN:CAA97460.1; P
R.Boy-Marcotte, E.; Danak, F.; Camonis, J.; Garreau, H.; Jacquet, M.
Gene 77, 21-30, 1989
Article: The C-terminal part of a gene partially homologous to CDC25 gene suppresses the
A.Reference number: PS0041; MUID:89306677; PMID:2545538
A.Accession: PS0041
A.Molecule type: DNA

A;Residues: 1-255 <BOY>
A;Cross-references: UNIPARC:UP100001746CF
A;Note: the authors translated the codon CAG for residue 248 as His
R;Purnelle, B.; Goffeau, A.
submitted to the EMBL Data Library, April 1996
A;Description: The sequence of 32 kb on the left arm of yeast chromosome XII reveals 14
mily and a new ABC transporter homologous to the human multidrug resistance protein.
A;Reference number: S69380
A;Accession: S69391
A;Molecule type: DNA
A;Residues: 1-1559 <PDR>
A;Cross-references: UNIPARC:UP10000126A9C; EMBL:X97560; NID:g1297003; PIDN:CAA62776.1; F
R;Mio99a, T.; Zimmermann, F.K.
Yeast 12, 693-708, 1996
A;Title: Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on a 43.7 kb
e conductance regulator protein CTR.
A;Reference number: S70557; MUID:96405918; PMID:8810043
A;Accession: S70560
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1559 <MIM>
A;Cross-references: UNIPARC:UP10000126A9C; EMBL:X91488; NID:g1495203; PIDN:CAA62776.1; F
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995
C;Genetics:
A;Gene: SGD:BPT1; MIPS:YLL015W
A;Cross-references: SGD:S0003938
A;Map position: 12L
C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C;Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F;28-44/Domain: transmembrane #status predicted <TM1>
F;143-159/Domain: transmembrane #status predicted <TM2>
F;178-194/Domain: transmembrane #status predicted <TM3>
F;334-350/Domain: transmembrane #status predicted <TM4>
F;421-437/Domain: transmembrane #status predicted <TM5>
F;526-542/Domain: transmembrane #status predicted <TM6>
F;550-566/Domain: transmembrane #status predicted <TM7>
F;654-647/Domain: ATP-binding cassette homology <ABC1>
F;672-679/Region: nucleotide-binding motif A (P-loop)
F;974-990/Domain: transmembrane #status predicted <TM8>
F;1017-1033/Domain: transmembrane #status predicted <TM9>
F;1099-1115/Domain: transmembrane #status predicted <TM10>
F;1118-1134/Domain: transmembrane #status predicted <TM11>
F;1212-1228/Domain: transmembrane #status predicted <TM12>
F;1319-1529/Domain: ATP-binding cassette homology <ABC2>
F;1336-1343/Region: nucleotide-binding motif A (P-loop)

Query Match 22.9%; Score 2294.5; DB 1; Length 1559;
Best Local Similarity 35.1%; Pred. No. 2.6e-146;
Matches 574; Conservative 292; Mismatches 541; Indels 227; Gaps 39;

QY 23 NTSPDTCFQNTVAVVPCFYIM-ACFPFYILSRHDSRYIQMTPLNKTATGFL- 80
DB 23 NALNP-----CFIVISAMQAVFLLIGSYQLMKLYKNNKVPKPPKPNFTLPSKINSRLT 78
QY 81 -LMIWCV-----ADLFYSFMRSGCFLAPVLPVLTLLG---TTLTAT 121
DB 79 HLTNVCQSTLILICELALVSSSDRYVPTLK-KALYTLNLF-----NLGSLFTQYLA 132
QY 122 FLIQLERRKGVSSGIMLTFWLVALVCALILRSKI-----MTALKEDAV 167
DB 133 F-----KSTFSWGNGLFYMPQILQLFLILORYHSSNRLTVISQTMILLEVL 185
QY 168 DLPRDITFYVFSLLIQVLVSCSDSPLEFSETHDPNCPRESSASFLRITFWITGL 227
DB 186 -LNSVAIFLY-DLCIFEP-----NELSEYKKGWY-----PVHLSYIFPIWNL 233
QY 228 IYVGYRPLESGDLMSLNK--EDTSEGVVPLVGNMKKECKTKRKQVKKVYSSKQDAP 285
DB 234 IVEYR-----NKIKDPNOPLPVPDLNLT----- 259
QY 286 KESSKYDANEVEALIVKSPQKWNPSLKFVLYKTFPPYFLMSFFFAHIDLMMFSGPQI 345
DB 260 SISKEFANWELKMLNRN-----SLMRAIWKSFRTISVAMLVETSDLSVQPOF 312

QY 346 IKLLIKFVN---DTKAPDMQGYFTVLLFVTACQLTVLHQYFHCIVSGMRKRTAVIGA 402
DB 313 IRIPTDGLNPTSSKYPPLNGVFATLFLVYISVSVLTNOFYGIEPAGIGISLASL 372
QY 403 VYRKALVITSARKSTVGEIVNLMSYDA---QRFMOLAATYINMWSAPLOVIALYLLW 459
DB 373 VYQSRLTLAERNKSTGDIINLMSVDVARIQGFENNA---QTIIAPIDIIYVLSLY 429
QY 460 LNIQPSVLAVAVVAVLVAVPVAVAMAKTKTYQVAHMSKONRIKLAMEIINGIVLYLYA 519
DB 430 WLGEKAVIGGLVTMAIMMPINAFISRYKKSLSTQMKYKDMIRITIELNAIKSIXLYA 489
QY 520 MELAFKQVLAIRE-ELKYLKKSAYLSAVNGTFVWCTPPLVNLACTRAVVTIDENNILD 578
DB 490 WEEMPMARLHVNRDMELKNFRKGIYSNLIYFAMNCVPLMWTCTSTGLP-SLPSDPLS 548
QY 579 AQTAFVSLATFNILRPLNLTLPWISISIVQASVSLKRLRFLSHELEPDSIER-RPVKD 637
DB 549 PAIVFPLSLFNILNSAIYVSPMINITIFTSVSMERLKGFLSDLEIDSTIERIDISAD 608
QY 638 GGGTNSITVRNATFTW-----ARSDPPLNGIT-----FSIPEGALV 674
DB 609 ERALPAIEMANNITFLMKSKSVLTSSQSGDMIRTDSESIIGSSQIALKNIDHFEAKRGDLV 668
QY 675 AVYQVCCGKSSLSALLAEMDKVEGH-----VAIKSVAVVPOQAMTQNDSLAENI 726
DB 669 CVGRVAGKSTFLKALIGQLPCWSSGRDSIPKLLIRSSVAACSGESIMWASVRENI 728
QY 727 LFGQLEPEPYRSYQACALLPDLTLPSPDRTEIGEGVNLSCGQKQRYSLARVYNSA 786
DB 729 LFGHKFPQDYDILTKACQLLPDLKILPDGDELTVGSKGISLGGQAKRLSLARVYSRA 788
QY 787 DIYLFDDPLSAVDAVAKHIFENV-IGPKMKLNKTRILVTHSMYSYLPQVDVITVMSGK 845
DB 789 DIYLLDDIISAVDAVEGKNIIEVYLIGKTKALNKNKTIILTTNVTYSLKHSQMIVALENGE 848
QY 846 ISEKSGYQELIAR---DGAFAEPLRTYAS---TEQDAEENGVTVSGPGKEAKOME 897
DB 849 IVEQGNVEDVMNRKNNTSKLKLLEEPDSIDNGNESDVOTERHSESEVEEP----- 900
QY 898 NGMLVDSAGKOROLSSSSSGDLSRHNSYAELOKAEKKEBTWKL----- 947
DB 901 -----LQKTYBETDEVT---EBBLIKANSRSLATLRPPFVGAOL 945
QY 948 -----MEADKAQTQGVKLSYVWDYWKAGLFIISLIFLM-CHNVSLASNYWLSLW 999
DB 946 DSVKTAQKAKETBEVGRVTKIYLAIVYKAGV-LGVVLFFLPHILTRVDFLAENFMLKYV 1004
QY 1000 TDDPIVNGOEHTRKRLSVYGAIGISQGIAPFGYMAVSI---GGLASRCLHYDLHSTI 1056
DB 1005 SEBENKSGNERWMMFVGVSLIGVAS--AAFNILRSIMMLVLCISIRGSKLHESMAKSV 1062
QY 1057 LRSPMSFEETPSGNLVNRSKELDVDSMIPIRYIKFMQSLFNVGACIVILLARTIAA 1116
DB 1063 IRSPMTFEETTPVGRILINRSSDMDAVDSMLQIYIFSEFFSILITLVTLVLGVNMBFL 1122
QY 1117 IILPGLIYFVQRFVAVASRQLEKLESVSRPVSHPNETLIGVSVIRAFEROFERFIH 1176
DB 1123 VFMNPLVLIYIYQTFYIVLSRELKRLISISYSPIMSLMSESLNGSIIIDAYHFERFIY 1182
QY 1177 QSDLKVDENQKAYVPSIVANRMLAVRECYGNCIYLPALPAVY---SRHSLSAGLVGLS 1233
DB 1183 LNYEKLQYVNDVFNFRSTNRKWSVRLQTIQATIVATATIALATNMVTKRQLSSGVGL 1242
QY 1234 VSYSLQVTVLNLVMSSEMETNIVAVRELKYSSEBEKAQPMQIETAPSPMPQVGRV 1293
DB 1243 MYSISLEVYGLTIVTTVTYIETNIVSVEHIVYCELPPEAGSINBEKRPDEMPKSGGI 1302
QY 1294 EFRNYCLAREEDLDFVLRHINVTINGEKVGIKRTGAGKSJTLGLFRINSABEGRII 1353
DB 1303 EFRNYSTKRENDPVLNINIVKIEPCEKGIYGRIGAGKSTLSLALFRILEPTEGKII 1362

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:56:46 ; Search time 245.527 Seconds
(without alignments)
5594.767 Million cell updates/sec

Title: US-10-665-283-8

Perfect score: 10016

Sequence: 1 MALRGFCASADSPDPLMDMNV.....RSVAVAKAKPKFSISPSLS 1947

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5.

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	78.5	1531	1 MRPI_HUMAN	P33527 homo sapien
2	7769	77.6	1515	2 O9UQ39_HUMAN	O9UQ39 homo sapien
3	7695	76.8	1531	2 O864R9_MACFA	O864R9 macaca fasc
4	7689	76.8	1531	2 O864S0_MACFA	O864S0 macaca fasc
5	7444	74.3	1459	2 O9UQ97_HUMAN	O9UQ97 homo sapien
6	7419.5	74.1	1456	2 O9UQAO_HUMAN	O9UQAO homo sapien
7	7272	72.6	1531	2 O6UR05_CANFA	O6UR05 canis famil
8	7242	72.3	1439	2 O59G19_HUMAN	O59G19 homo sapien
9	7230.5	72.2	1530	2 O8HX05_BOVIN	O8HX05 bos taurus
10	7104.5	70.9	1400	2 O9UQ98_HUMAN	O9UQ98 homo sapien
11	7002.5	69.9	1528	1 MRPI_MOUSE	O35379 mus musculu
12	6932.5	69.2	1532	2 O810E4_RAT	O810E4 rattus norv
13	6932.5	69.2	1532	2 O8CG09_RAT	O8CG09 rattus norv
14	6892	68.8	1523	2 O810G9_RAT	O810G9 rattus norv
15	6301	62.9	1523	2 O5F3E4_CHICK	O5F3E4 gallus gall
16	6075	60.7	1215	2 O68CP7_HUMAN	O68CP7 homo sapien
17	4526	45.2	1522	2 O5SUF4_MOUSE	O5SUF4 mus musculu
18	4522	45.1	1519	2 O80ZK8_MOUSE	O80ZK8 mus musculu
19	4522	45.1	1522	2 O56PH0_MOUSE	O56PH0 mus musculu
20	4508.5	45.0	1523	2 O59DU0_MOUSE	O59DU0 mus musculu
21	4481.5	44.8	1527	1 MRPI_HUMAN	O51438 homo sapien
22	4481.5	44.8	1533	1 O59H05_HUMAN	O59H05 homo sapien
23	4422	44.1	1514	2 O96Q09_HUMAN	O96Q09 homo sapien
24	4402	43.9	1498	2 O59DK9_MOUSE	O59DK9 mus musculu
25	4396	43.9	1522	1 MRPI_RAT	O85573 rattus norv
26	4148.5	41.4	1496	2 O4S7E7_TETNG	O4S7E7 tetradon n
27	3885.5	38.8	1548	2 O7KTC3_DROME	O7KTC3 drosophila
28	3880.5	38.7	1548	2 O7KTC1_DROME	O7KTC1 drosophila
29	3879.5	38.7	1548	2 O7KTC2_DROME	O7KTC2 drosophila
30	3874.5	38.7	1548	2 O7KTC2_DROME	O7KTC2 drosophila
31	3860.5	38.5	1548	2 O917NO_DROME	O917NO drosophila

32	3857	38.5	1549	2 O7KTC0_DROME	O7KTC0 drosophila
33	3845.5	38.4	1548	2 O8T9C5_DROME	O8T9C5 drosophila
34	3838.5	38.3	1548	2 O7KTB8_DROME	O7KTB8 drosophila
35	3836.5	38.3	1548	2 O7KTB9_DROME	O7KTB9 drosophila
36	3818.5	38.1	1548	2 O7KTD0_DROME	O7KTD0 drosophila
37	3813.5	38.1	1548	2 O7KTC8_DROME	O7KTC8 drosophila
38	3812.5	38.1	1548	2 O7KTC4_DROME	O7KTC4 drosophila
39	3807.5	38.0	1548	2 O7KTC9_DROME	O7KTC9 drosophila
40	3793.5	37.9	1548	2 O7KTC7_DROME	O7KTC7 drosophila
41	3790	37.8	1549	2 O9VK56_DROME	O9VK56 drosophila
42	3771.5	37.7	1548	2 O7KTC5_DROME	O7KTC5 drosophila
43	3769.5	37.6	1548	2 O7KTC6_DROME	O7KTC6 drosophila
44	3741	37.4	1564	2 O8QC98_RAJER	O8QC98 raja erinac
45	3738.5	37.3	1567	2 O6PH26_BRARE	O6PH26 brachydanio

ALIGNMENTS

RESULT 1
MRPI_HUMAN STANDARD; PRT; 1531 AA.
ID MRPI_HUMAN
AC P33527; O14819; P78419;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).
GN Name=ABCC1; Synonyms=MRP, MRP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN MRPI_HUMAN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93088080; PubMed=1360704;
RA Cole S.P.C., Bhargava G., Gerlach J.H., Mackie J.E., Grant C.E., Almqvist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;
RT "Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line.";
RL Science 258:1650-1654 (1992).
[2]
RN MRPI_HUMAN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93262415; PubMed=8098549;
RA Cole S.P.C., Deeley R.G.;
RT "Multidrug resistance-associated protein: sequence correction.";
RL Science 260:879-879 (1993).
[3]
RN MRPI_HUMAN
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96251691; PubMed=8649356;
RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;
RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter.";
RL Mol. Pharmacol. 49:962-971 (1996).
[4]
RN MRPI_HUMAN
RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1999.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378 (1997).
[5]
RN MRPI_HUMAN
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA] OF 1131-1531.
RX MEDLINE=99425270; PubMed=10493829; DOI=10.1006/geno.1999.5927;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhrmann U., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";
RL Genomics 60:295-308 (1999).

[6]
 RP TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=97444425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;
 RA Hifner D.R., Almquist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,
 RA Deley R.G., Cole S.P.C.,
 RT "Membrane topology of the multidrug resistance protein (MRP). A study
 of glycosylation-site mutants reveals an extracytosolic NH2
 terminus.";
 RT J. Biol. Chem. 272:23623-23630 (1997).
 RN [7]
 RP TOPOLOGY.
 RX MEDLINE=97476249; PubMed=9334225; DOI=10.1074/jbc.272.42.26479;
 RA Kaet C., Gros P.;
 RT "Topology mapping of the amino-terminal half of multidrug resistance-
 RT associated protein by epitope insertion and immunofluorescence.";
 RL J. Biol. Chem. 272:26479-26487 (1997).
 RN [8]
 RP TOPOLOGY.
 RX MEDLINE=96153110; PubMed=9485377; DOI=10.1021/bi972332v;
 RA Kaet C., Gros P.;
 RT "Epitope insertion favors a six transmembrane domain model for the
 RT carboxy-terminal portion of the multidrug resistance-associated
 RT protein.";
 RL Biochemistry 37:2305-2313 (1998).
 RN [9]
 RP MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-1455.
 RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abdi.2001.2441;
 RA Cui L., Hou Y.-X., Riordan J.R., Chang X.-B.;
 RT "Mutations of the Walker B motif in the first nucleotide binding
 RT domain of multidrug resistance protein MRP1 prevent conformational
 RT maturation.";
 RL Arch. Biochem. Biophys. 392:153-161 (2001).
 RN [10]
 RP MUTAGENESIS OF TRP-1246.
 RX MEDLINE=21238219; PubMed=11278867; DOI=10.1074/jbc.M011246200;
 RA Ito K., Olsen S.L., Qiu W., Deeley R.G., Cole S.P.C.;
 RT "Mutation of a single conserved tryptophan in multidrug resistance
 RT protein 1 (MRP1/ABCC1) results in loss of drug resistance and
 RT selective loss of organic anion transport.";
 RL J. Biol. Chem. 276:15616-15624 (2001).
 RN [11]
 RP VARIANTS GLN-633 AND VAL-671.
 RX MEDLINE=2029630; PubMed=10835642; DOI=10.1038/76102;
 RA Le Saux O., Urban Z., Tschuch C., Galsazar K., Bacchelli B.,
 RA Ougelli O., Pasquali-Ronchetti I., Pope F.M., Richards A., Terry S.,
 RA Becroft L., de Paep A., Boyd C.D.;
 RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma
 RT elasticum.";
 RL Nat. Genet. 25:223-227 (2000).
 RN [12]
 RP VARIANT VAL-671.
 RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;
 RA Ringsfeil F., Lehwahl M.G., Cristiano A.M., Utito J.;
 RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a
 RT transmembrane ATP-binding cassette (ABC) transporter.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006 (2000).
 RN [13]
 RP VARIANT SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.
 RX MEDLINE=21576494; PubMed=11721885; DOI=10.1007/s100380170017;
 RA Conrad S., Kaufmann H.-M., Ito K., Deeley R.G., Cole S.P.C.,
 RA Schrenk D.;
 RT "Identification of human multidrug resistance protein 1 (MRP1)
 RT mutations and characterization of a G671V substitution.";
 RL J. Hum. Genet. 46:656-663 (2001).
 RN [14]
 RP VARIANTS MET-117 AND LEU-1512.
 RX MEDLINE=20579883; PubMed=11139250;
 DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14>3.0.CO;2-F;
 RA Perdu J., Germain D.P.;
 RT "Identification of novel polymorphisms in the pms and MRP1 (ABCC1)
 RT genes at locus 16p13.1 and exclusion of both genes as responsible for
 RT pseudoxanthoma elasticum.";
 RL Hum. Mutat. 17:74-75 (2001).

[15]
 RP VARIANTS SER-43; ILE-73; GLN-723 AND GLN-1058.
 RX MEDLINE=21163848; PubMed=11266082;
 DOI=10.1097/00008571-200103000-00008;
 RA Ito S., Ieiri I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;
 RT "Polymorphism of the ABC transporter genes, MRP1, MRP2 and MRP2/CMOAT,
 RT in healthy Japanese subjects.";
 RL Pharmacogenetics 11:175-184 (2001).
 CC -1- FUNCTION: May participate directly in the active transport of
 CC drugs into subcellular organelles or influence drug distribution
 CC indirectly.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=8;
 CC Comment-Additional isoforms seem to exist. Experimental
 CC confirmation may be lacking for some isoforms;
 CC Name=Allexons;
 CC IsoId=P33527-1; Sequence=Displayed;
 CC Name=Delexon-17;
 CC IsoId=P33527-2; Sequence=VSP_000037;
 CC Name=Delexon-18;
 CC IsoId=P33527-3; Sequence=VSP_000038;
 CC Name=Delexon-30;
 CC IsoId=P33527-4; Sequence=VSP_000039;
 CC Name=Delexon-17-18;
 CC IsoId=P33527-5; Sequence=VSP_000037, VSP_000038;
 CC Name=Delexon-17-30;
 CC IsoId=P33527-6; Sequence=VSP_000037, VSP_000039;
 CC Name=Delexon-18-30;
 CC IsoId=P33527-7; Sequence=VSP_000038, VSP_000039;
 CC Name=Delexon-17-18-30;
 CC IsoId=P33527-8; Sequence=VSP_000037, VSP_000038, VSP_000039;
 CC -1- TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear
 CC cells.
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
 CC -1- SIMILARITY: Contains 2 ABC transporter domains.
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.inbio.gen.fr/services/chromocancer/gene/MRP106.html".
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: L05628; AAB46616.1; -, mRNA.
 CC EMBL: AF022853; AAB83983.1; -, Genomic DNA.
 CC EMBL: AF022824; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022825; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022826; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022827; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022828; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022829; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022830; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022831; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022832; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022833; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022834; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022835; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022836; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022837; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022838; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022839; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022840; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022841; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022842; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022843; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022844; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022845; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022846; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022847; AAB83983.1; JOINED; Genomic DNA.
 CC EMBL: AF022848; AAB83983.1; JOINED; Genomic DNA.

DR EMBL; AF022849; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83983.1; JOINED; Genomic DNA.

Query Match 78.5%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALRGFCSDGSDPLDMNMTNTSNDPFCFQNTLVNWPCEYLACPFEPFLYSRH 60
DB 1 MALRGFCSDGSDPLDMNMTNTSNDPFCFQNTLVNWPCEYLACPFEPFLYSRH 60
QY 61 DRGYIQTPLNKTALGFLMIYCMADLFYSFWEKRGIFLAFVLSPTLIGITTLA 120
DB 61 DRGYIQTPLNKTALGFLMIYCMADLFYSFWEKRGIFLAFVLSPTLIGITTLA 120
QY 121 TPLIQLERRKGVSSGIMLTFMLVAVCALAIRSKMTALKEDAOVDLPFDITFYVYS 180
DB 121 TPLIQLERRKGVSSGIMLTFMLVAVCALAIRSKMTALKEDAOVDLPFDITFYVYS 180
QY 181 LLLIQLVLSGFSRSPLESETIHDNPNCPSSSASFSLRITFWITGLIVRGYRQLSGSD 240
DB 181 LLLIQLVLSGFSRSPLESETIHDNPNCPSSSASFSLRITFWITGLIVRGYRQLSGSD 240
QY 241 LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVKVSXDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVKVSXDPAPKESKVDANEVEAL 300
QY 301 IVSPKQKPNPSLFKVLKTFPGPYFLMSFPFKAIHDLMSFGPOILKLIKFNVDYAPD 360
DB 301 IVSPKQKPNPSLFKVLKTFPGPYFLMSFPFKAIHDLMSFGPOILKLIKFNVDYAPD 360
QY 361 MOGYFYVLLFPAACLOTVLHOYFHTCFVSGMRKTAIVGAYVRKALVITNSARKSSTV 420
DB 361 MOGYFYVLLFPAACLOTVLHOYFHTCFVSGMRKTAIVGAYVRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMISAPLOVIALYLLMLNGSVLAGVAVMLAMPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMISAPLOVIALYLLMLNGSVLAGVAVMLAMPVN 480
QY 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILINGIKVLYAMELAFKDYAIRBELKVLK 540
DB 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILINGIKVLYAMELAFKDYAIRBELKVLK 540
QY 541 KSAVLSAVGFTWCPPELVALCTPAVYVITDENNITADTAFAFSLAFMLPPTNLP 600
DB 541 KSAVLSAVGFTWCPPELVALCTPAVYVITDENNITADTAFAFSLAFMLPPTNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEBPSIERRPVKDGGTNSITVYNAFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEBPSIERRPVKDGGTNSITVYNAFTWARSDDPT 660
QY 661 LMGITTSIPFGALVAVVGQVCGCKSSLALALEMDKVEGHVAIKGSVAVYPOQAMTQND 720
DB 661 LMGITTSIPFGALVAVVGQVCGCKSSLALALEMDKVEGHVAIKGSVAVYPOQAMTQND 720
QY 721 SLRENTLFGQLEEPYRSYIOACALLPDLIELIPSGRTIIGKGVNLSSGQQRVSILAR 780
DB 721 SLRENTLFGQLEEPYRSYIOACALLPDLIELIPSGRTIIGKGVNLSSGQQRVSILAR 780
QY 781 AAVSNADIVLPDPLSAVDHVGHIIPENYIGPKMKNTRIILVTHSMGYLPQVDYIV 840
DB 781 AAVSNADIVLPDPLSAVDHVGHIIPENYIGPKMKNTRIILVTHSMGYLPQVDYIV 840
QY 841 MSGGKISEMGSYELLARDGAFAPFLTYASTROEDDAENGVTVSGPGEAKAKOMENGM 900
DB 841 MSGGKISEMGSYELLARDGAFAPFLTYASTROEDDAENGVTVSGPGEAKAKOMENGM 900
QY 901 LVYDSAGKOLQRLSSSSSGDISRHNSATIELQKAEKPEWKLMEADKQOTGVKL 960
DB 901 LVYDSAGKOLQRLSSSSSGDISRHNSATIELQKAEKPEWKLMEADKQOTGVKL 960
QY 961 SVYWDYKAIQGLIFSLIFLFCMNHVSALASNYMLSLMTDDPIVNGTOHTKYRLSVYG 1020
DB 961 SVYWDYKAIQGLIFSLIFLFCMNHVSALASNYMLSLMTDDPIVNGTOHTKYRLSVYG 1020
```

```
DB 961 SVYWDYKAIQGLIFSLIFLFCMNHVSALASNYMLSLMTDDPIVNGTOHTKYRLSVYG 1020
QY 1021 ALGISGIAVFGYMAVSTIGIILASRCLHVDLHSLIRSPMSFEFRPSGNLVNRSKEL 1080
DB 1021 ALGISGIAVFGYMAVSTIGIILASRCLHVDLHSLIRSPMSFEFRPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPETVYKMFMSLFWVIGACIVYLLATPILAIIPPLGIYFFVORFYVASSROL 1140
DB 1081 DTVDSMIPETVYKMFMSLFWVIGACIVYLLATPILAIIPPLGIYFFVORFYVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQAXYPSIVANWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQAXYPSIVANWLA 1200
QY 1201 VRLCEVNCIVLPAALFAVISRHSLSAGLVGSVSLQVTTIYMLNLRSSSEMETIYA 1260
DB 1201 VRLCEVNCIVLPAALFAVISRHSLSAGLVGSVSLQVTTIYMLNLRSSSEMETIYA 1260
QY 1261 VERLKEYSETEKAPMOIQETAPSSWPQVGRVFPNRYCLRYREDLPVLRHINVTING 1320
DB 1261 VERLKEYSETEKAPMOIQETAPSSWPQVGRVFPNRYCLRYREDLPVLRHINVTING 1320
QY 1321 EKVGIYGRGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKITTIIPODPVL 1380
DB 1321 EKVGIYGRGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKITTIIPODPVL 1380
QY 1381 SGSLRNWLDPPSGYSBEEVTSLELAHLKDFVALPDKLDHECAEGGENISVQORQVCL 1440
DB 1381 SGSLRNWLDPPSGYSBEEVTSLELAHLKDFVALPDKLDHECAEGGENISVQORQVCL 1440
QY 1441 ARALLKTKTLVDEADTAANDLETDDLIQSTIRFOEDCTVLTARHANTIMYTRYIVL 1500
DB 1441 ARALLKTKTLVDEADTAANDLETDDLIQSTIRFOEDCTVLTARHANTIMYTRYIVL 1500
QY 1501 DRKEIQEYGAPSDLLQORGLFYSMAKDAGIV 1531
DB 1501 DRKEIQEYGAPSDLLQORGLFYSMAKDAGIV 1531

RESULT 2
Q9U099 HUMAN
ID Q9U099 HUMAN PRELIMINARY; PRT; 1515 AA.
AC Q9U099;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT Analysis of the inticon-exon organization of the human multidrug-
  resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC - - Similarity: Belongs to the ABC transporter family.
DR EMBL; AF022853; AAB83979.1; -; Genomic DNA.
DR EMBL; AF022827; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022828; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022829; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022831; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022839; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022849; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022849; AAB83979.1; JOINED; Genomic DNA.
```

DR	EML1	AF022848	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022847	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022846	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022845	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022844	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022843	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022842	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022852	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022851	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022840	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022838	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022836	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022834	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022832	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022826	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022825	AAB83979.1	JOINED	Genomic DNA
DR	EML1	AF022830	AAB83979.1	JOINED	Genomic DNA
DR	HSSP	P08716	IMT0		
DR	Ensembl	ENSG00000103222	Homo sapiens		
DR	GO	GO:0016021	C:integral to membrane		IEA.
DR	GO	GO:0005524	F:ATP binding		IEA.
DR	GO	GO:0016882	F:ATPase activity		IEA.
DR	GO	GO:0042656	F:ATPase activity, coupled to transmembrane m. .		IEA.
DR	GO	GO:0000166	F:nucleotide binding		IEA.
DR	GO	GO:0006810	P:transport		IEA.
DR	InterPro	IPR003593	AAA ATPase		
DR	InterPro	IPR011527	ABC transporter_1		
DR	InterPro	IPR001140	ABC_TM_transp.		
DR	InterPro	IPR003439	ABC_transp_like		
DR	InterPro	IPR005292	MRF_aescoc		
DR	InterPro	IPR007119	Prot_kinase		
DR	Pfam	PF00664	ABC_membrane1_2		
DR	Pfam	PF00005	ABC_tran_2		
DR	PfDom	PD000006	ABC_transporter; 2		
DR	SMART	SM00382	AAA_2		
DR	TIGRFAMs	TIGR00957	MRF_aescoc_pro; 1		
DR	PROSITE	PS50929	ABC_TM1F_2		
DR	PROSITE	PS00211	ABC_TRANSPORTER_1; 2		
DR	PROSITE	PS50893	ABC_TRANSPORTER_2; 2		
DR	PROSITE	PS00107	PROTEIN_KINASE_ATP; UNKNOWN_1		
KM	ATP-binding	Nucleotide-binding			
FT	NON_TER	1			
SQ	SEQUENCE	1515 AA; 169853 MW; DC85592817C439PE CRC64;			
<hr/>					
Query Match		77.6%; Score 7769; DB 2; Length 1515;			
Best Local Similarity		100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0			
Matches 1515; Conservative		0; Mismatches 0; Indels 0; Gaps 0			
OY	17	DNNVTWNTSNPDFTRKCFONTVLVWPCCFLTMACFPFFPYLISRHRDGYIOMTPLNKTXTA	76		
DB	1	DNNVTWNTSNPDFTRKCFONTVLVWPCCFLTMACFPFFPYLISRHRDGYIOMTPLNKTXTA	60		
OY	77	LGFLIMTYCMADLFSPWERSRGIFPLAPFYLVNSPTILGTTTLIAATPIOLERRKGVSQG	136		
DB	61	LGFLIMTYCMADLFSPWERSRGIFLAPFYLVNSPTILGTTTLIAATPIOLERRKGVSQG	120		
OY	137	IMLTFEMLVALVCALAILRSKITAKLEDAQVDLFSDITFYYVFSLILLIQLVLSCSFSDSRP	196		
DB	121	IMLTFEMLVALVCALAILRSKITAKLEDAQVDLFSDITFYYVFSLILLIQLVLSCSFSDSRP	180		
OY	197	LFSETHDPNCPRESSASFLSRITTWMTITGLIVRGYROPLEGSDIMSLINKEDTSSQVVAV	256		
DB	181	LFSETHDPNCPRESSASFLSRITTWMTITGLIVRGYROPLEGSDIMSLINKEDTSSQVVAV	240		
OY	257	LYKNMKKKCAKTRKOPVKVNVYSXKPOAPCKESSKYVDANEVEVALIVKSQCKEMNSLPFY	316		
DB	241	LYKNMKKKCAKTRKOPVKVNVYSXKPOAPCKESSKYVDANEVEVALIVKSQCKEMNSLPFY	300		
OY	317	LYKTGPAPFLMSFFPKAIHDLMMFGSQPOLIKLLIKFVNDTKA.PDMQGYFYTVLLEFYTA	376		
DB	301	LYKTGPAPFLMSFFPKAIHDLMMFGSQPOLIKLLIKFVNDTKA.PDMQGYFYTVLLEFYTA	360		

QY	377	QTLVHQQYEHICFVSGMRKTAIVGAYRRALVITNSARKSSTGELIVNLSVDAOEPM	436
Db	361	QTLVHQYFHHCFVSGMRKTAIVGAYRRALVITNSARKSSTGELIVNLSVDAOEPM	420
QY	437	LATVINNMWSPLOVILALVYLMNLCPSVLAGAVAVWVLMPPVNAVAMAKTKYQVAHMK	496
Db	421	LATVINNMWSPLOVILALVYLMNLCPSVLAGAVAVWVLMPPVNAVAMAKTKYQVAHMK	480
QY	497	SKDNRIKLMNEILNGIKVLKLYAMELAPKOKVLAIRQEBKVLKKSAYLSAVGTFWVCT	556
Db	481	SKDNRIKLMNEILNGIKVLKLYAMELAPKOKVLAIRQEBKVLKKSAYLSAVGTFWVCT	540
QY	557	PELVALCFPAVYVYVTDENNIIIDACTAVSVALFPHILAPLNTLPMVTSSTVQASVSKRL	616
Db	541	PELVALCFPAVYVYVTDENNIIIDACTAVSVALFPHILAPLNTLPMVTSSTVQASVSKRL	600
QY	617	RIPLSHSELEPDSIERRPVKDGGTNSITVYNAFTFWARSDPPTLNGITFSIPEGALYAV	676
Db	601	RIPLSHSELEPDSIERRPVKDGGTNSITVYNAFTFWARSDPPTLNGITFSIPEGALYAV	660
QY	677	VGQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVYPOQAMIQNDSLRENTILFGCOLEBPY	736
Db	661	VGQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVYPOQAMIQNDSLRENTILFGCOLEBPY	720
QY	737	YRSYVIOACALLPDEILIPSGDRTEIGEGKVNLSGQOKORVSLARAYVSNADITYFDDPLS	796
Db	721	YRSYVIOACALLPDEILIPSGDRTEIGEGKVNLSGQOKORVSLARAYVSNADITYFDDPLS	780
QY	797	AVDAHVGHIEPENTYIGPKMLKNTKRIILVTHSMGYLPOVDVITYMSGGKISEMGSYQEBL	856
Db	781	AVDAHVGHIEPENTYIGPKMLKNTKRIILVTHSMGYLPOVDVITYMSGGKISEMGSYQEBL	840
QY	857	ARDGAFAEFLTYASTEOEDAEENGVTGVSGBPKEAKOMENGMVLTDSACKOLQOROLSS	916
Db	841	ARDGAFAEFLTYASTEOEDAEENGVTGVSGBPKEAKOMENGMVLTDSACKOLQOROLSS	900
QY	917	SSSYSGDISRRHNSTAELOKAKAKKETWKMEDADKAQTQYKLSYVYDNYKAIGLPISF	976
Db	901	SSSYSGDISRRHNSTAELOKAKAKKETWKMEDADKAQTQYKLSYVYDNYKAIGLPISF	960
QY	977	LSIFLPMONNHSALASNYWLSLMTDDPIVNGTOHKTVRLSVYGALGISOGIAVFGYSMA	1036
Db	961	LSIFLPMONNHSALASNYWLSLMTDDPIVNGTOHKTVRLSVYGALGISOGIAVFGYSMA	1020
QY	1037	VSIGGILASRCIAHYDLHLSILRSPMSFEERPPSGNLVNRFSKEJLDTVDSMIDEVIKMPMG	1096
Db	1021	VSIGGILASRCIAHYDLHLSILRSPMSFEERPPSGNLVNRFSKEJLDTVDSMIDEVIKMPMG	1080
QY	1097	SLFNVYIGACIYILLATPIAIIIPPLGIITYFVORFYVASSRQOLKRLSVSRSPYSHFN	1156
Db	1081	SLFNVYIGACIYILLATPIAIIIPPLGIITYFVORFYVASSRQOLKRLSVSRSPYSHFN	1140
QY	1157	ETLIGVSYTRAFEEOEERFIHOSDLKDNENQAAVYPSIVANWMLVRLCEVNCICVLFPAL	1216
Db	1141	ETLIGVSYTRAFEEOEERFIHOSDLKDNENQAAVYPSIVANWMLVRLCEVNCICVLFPAL	1200
QY	1217	FAVISRHSLSAGVLGVLSVSYSLQVTTYLNMVLRMSSEMETNIVAVERLKEYSETEKEAPM	1276
Db	1201	FAVISRHSLSAGVLGVLSVSYSLQVTTYLNMVLRMSSEMETNIVAVERLKEYSETEKEAPM	1260
QY	1277	QIOETAPPSSWPQVGRVBFNNYCLTRYEDDLFVLRHINVITNGEBKVGIVGRTGAGKSSL	1336
Db	1261	QIOETAPPSSWPQVGRVBFNNYCLTRYEDDLFVLRHINVITNGEBKVGIVGRTGAGKSSL	1320
QY	1337	TLGFLFRINESAEGEIIIDGINIAKIGLHDLAFKTIIPQDVLPSGSLRMLDPSPSOYSD	1396
Db	1321	TLGFLFRINESAEGEIIIDGINIAKIGLHDLAFKTIIPQDVLPSGSLRMLDPSPSOYSD	1380
QY	1397	EEWMTSTELAHKOPVSALPDKLDBECABEGENISVGOROLVCLARALARTKIKLIVDEA	1456
Db	1381	EEWMTSTELAHKOPVSALPDKLDBECABEGENISVGOROLVCLARALARTKIKLIVDEA	1440
QY	1457	TPAADVLEFDDLIQSTIRIQPEDCIVLTJAHNLYNTIMDTRVIVLDKGEIOEYGAPSDILO	1516

|||||
Db 1441 TAAVDELTDLIGSTIRQFEDCTVLTARHLNTIMOTRYIVJDKEIGGYGAPSDLLQ 1500
QY 1517 ORGLFYSMAXDAGLV 1531
Db 1501 ORGLFYSMAXDAGLV 1515
RESULT 3
0864R9_MACFA
ID 0864R9_MACFA PRELIMINARY; PRT; 1531 AA.
AC 0864R9;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, last annotation update)
DE Multidrug resistance protein 1B.
GN Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22544876; PubMed=12657726;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Perry W.L., Godinot N.; EMBL/GenBank/DBJ databases.
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY146673; AAN65349.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042666; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; KRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR Prodom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; KRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF_2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1531 AA; 171659 MW; 1AE788EFD9EF459 CRC64;
Query Match 76.8%; Score 7695; DB 2; Length 1531;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 121 TFLIQLERRKGVSSGIMLTFWLVAVCALAIILRSKIMTALKEDAQVDLFRDITFVYYS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVAVCALAIILRSKIMTALKEDAQVDLFRDITFVYYS 180
QY 181 LLLIQLVLSGCRSDRPLPSETIHDNPDPCSSASFLSRIITFWITGILVGRYOPLEBSD 240
Db 181 LVLQLVLSGCRSDRPLPSETIHDNPDPCSSASFLSRIITFWITGILVGRYOPLEBSD 240
QY 241 LMSLNKEDTSBQVVPVYKWKWKGCACTRQKPVKVVYSSKDPAOPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSBQVVPVYKWKWKGCACTRQKPVKVVYSSKDPAOPKSSKVDANEVEAL 300
QY 301 IVKSPQKEMPSLPEKVLTKTGGPYFLMSFFPKAHDLMFSGPOLILLLKFNVDTPAD 360
Db 301 IVKSPQKEMPSLPEKVLTKTGGPYFLMSFFPKAHDLMFSGPOLILLLKFNVDTPAD 360
QY 361 MOGYFTVLLFVTAQLOTVLVHOYFHI CFVSGMRKTAIVIGAVYRKALVITNSARKSTV 420
Db 361 MOGYFTVLLFVTAQLOTVLVHOYFHI CFVSGMRKTAIVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMVSVDQRFMDLATTINMTWSAPLOYITLALYLMNLGPSVLAQVAVMTLMPVN 480
Db 421 GEIVNLMVSVDQRFMDLATTINMTWSAPLOYITLALYLMNLGPSVLAQVAVMTLMPVN 480
QY 481 AVYAMKTKTYQVAMHMSKDKNRKILMNLINGIKVLKIYANELAKDKVLAIROBELVYLK 540
Db 481 AVYAMKTKTYQVAMHMSKDKNRKILMNLINGIKVLKIYANELAKDKVLAIROBELVYLK 540
QY 541 KSAVLISAVGTFTWCTPFLVALCTFAVYVITIDENNILDAOTAEPFSLALFNILRPILNP 600
Db 541 KSAVLISAVGTFTWCTPFLVALCTFAVYVITIDENNILDAOTAEPFSLALFNILRPILNP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELBDSIERRPVKDGGTNSITVANATFTMARSDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEELBDSIERRPVKDGGTNSITVANATFTMARSDPT 660
QY 661 LINGTFSIPGALVAVVGVGCCGKSSLSLLALMEDVBEHVALKGSVAVPOQAWIQND 720
Db 661 LINGTFSIPGALVAVVGVGCCGKSSLSLLALMEDVBEHVALKGSVAVPOQAWIQND 720
QY 721 SLRENILFGQLEBPYRSYIOACALPDLEILPSCGRTIEGKVMISGGOKORVSLAR 780
Db 721 SLOENILFGQLEBPYRSYIOACALPDLEILPSCGRTIEGKVMISGGOKORVSLAR 780
QY 781 AVYSNADIYLPDPLSAVDAHVGHIPENVIIPKGMKNKTRILVTHSMSYLPQVDYIIV 840
Db 781 AVYSNADIYLPDPLSAVDAHVGHIPENVIIPKGMKNKTRILVTHSMSYLPQVDYIIV 840
QY 841 MSGKTISEMSYQELLARDGAPAFELTASTOEOAEENGVTGVSQPKKAKOMENGM 900
Db 841 MSGKTISEMSYQELLARDGAPAFELTASTOEOAEENGVTGVSQPKKAKOMENGM 900
QY 901 LVTDSAGKOROLSSSSSYSGDISRRHNSHTAELOKAKKEFTWKLMKADKAOYGOKL 960
Db 901 LVTDSAGKOROLSSSSSYSGDISRRHNSHTAELOKAKKEFTWKLMKADKAOYGOKL 960
QY 961 SYTWYDVKAKGLPFLSLFLFMCNHYASALASNYWLSLWTDPIVNGTOEHTKYRLSVYG 1020
Db 961 SYTWYDVKAKGLPFLSLFLFMCNHYASALASNYWLSLWTDPIVNGTOEHTKYRLSVYG 1020
QY 1021 ALGISQCIANFGYSMAVSTIGIILASRCLHVDLHSLIRSPSPERFPGSNLVNRSKEL 1080
Db 1021 ALGISQCIANFGYSMAVSTIGIILASRCLHVDLHSLIRSPSPERFPGSNLVNRSKEL 1080
QY 1081 DTVDSMIPVYIKPMFMSLFVNIAGACIYIILATPIAIIIPPLGIYFFVORFYAASROL 1140
Db 1081 DTVDSMIPVYIKPMFMSLFVNIAGACIYIILATPIAIIIPPLGIYFFVORFYAASROL 1140
QY 1141 KRLSVSRSGVYGHFNELLGVSVITRAFEBQERFIHOSDKVDENOKAYIPSYANRWLA 1200
Db 1141 KRLSVSRSGVYGHFNELLGVSVITRAFEBQERFIHOSDKVDENOKAYIPSYANRWLA 1200
QY 1201 VRLCVCNCLVFAALFAVYSRSLSGVLGVSYSLOQYTTIYLMNVRRSMENETIVA 1260

Db 1201 VRLCVCAGCIVLFAALFAVIRSHSISAGLVGSYSIQVTTYLNMVLVMSSEMETIVA 1260
QY 1261 VERLKEVSETEKEAPWQIOETAPPSWPQVGVVEFRNYCLARYEDLDVLRHINVTINGG 1320
Db 1261 VERLKEVSETEKEAPWQIOETAPPSWPQVGVVEFRNYCLARYEDLDVLRHINVTINGG 1320
QY 1321 EKVGVIGRTGAGKSSLTGLFRINSAAGEI11IDGINIAKIGLHDLRFKRTI11IPQDPVLF 1380
Db 1321 EKVGVIGRTGAGKSSLTGLFRINSAAGEI11IDGINIAKIGLHDLRFKRTI11IPQDPVLF 1380
QY 1381 SGSLRMLNDPPSQSDEEVMWLSLELAHKDFVSALPKDLHHCAGGEMLSVGQRLVCL 1440
Db 1381 SGSLRMLNDPPSQSDEEVMWLSLELAHKDFVSALPKDLHHCAGGEMLSVGQRLVCL 1440
QY 1441 ARALLRKRTIIVLDEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKRTIIVLDEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOEYGAPSDLLQQRGLFYSMKXAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQQRGLFYSMKXAGLV 1531

RESULT 4

Q864S0_MACFA PRELIMINARY; PR1; 1531 AA.
ID Q864S0; AC Q864S0;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Multidrug resistance protein 1A.
GN Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Perry W.L., Godinot N.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY146672; AA065346.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane.1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR00719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane.2.
DR Pfam; PD00005; ABC_tran.2.
DR Pfam; PD00006; ABC_transporter.2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.

DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SEQUENCE 1531 AA; 171701 MW; 168712B8CC2D2B89 CRC64;

Query Match 76.8%; Score 7689; DB 2; Length 1531;
Best local similarity 97.6%; Pred. No. 0;
Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

1 MALRGFCSADGSDLLMDMNTVMTNTPNDPFCRQNTLVWPCGYLWACFPFYLYSRH 60
1 MALRGFCSADGSDLLMDMNTVMTNTPNDPFCRQNTLVWPCGYLWACFPFYLYSRH 60
QY DRGYQMTPLNKTALGFLIMVCMADLFYSFMERSRGIFLAVFLVSPYLLGITTLA 120
Db DRGYQMTPLNKTALGFLIMVCMADLFYSFMERSRGIFLAVFLVSPYLLGITTLA 120
QY 61 DRGYQMTPLNKTALGFLIMVCMADLFYSFMERSRGIFLAVFLVSPYLLGITTLA 120
Db 61 DRGYQMTPLNKTALGFLIMVCMADLFYSFMERSRGIFLAVFLVSPYLLGITTLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVAVCALAIRSKIMTALKEBQVDFRFDITFYVYS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVAVCALAIRSKIMTALKEBQVDFRFDITFYVYS 180
QY 181 LLLIQVLVSCFSDSPFSEITIHDNPPCESSASFLSRIITFWITGLIVRGYRQPLEGSD 240
Db 181 LVLIQVLVSCFSDSPFSEITIHDNPPCESSASFLSRIITFWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLYKWKKECAKTRKQPVVYVSSKDPAPKESKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLYKWKKECAKTRKQPVVYVSSKDPAPKESKVDANEVEAL 300
QY 301 IVKSPQKEWNSLFKVLKTFEPYFLMSFPFKAIHDLMPESGPOILKLLIKFVNDTXAPD 360
Db 301 IVKSPQKEWNSLFKVLKTFEPYFLMSFPFKAIHDLMPESGPOILKLLIKFVNDTXAPD 360
QY 361 WQGFYTVLAEVTAACLOTIVLHOYFHI CFVSGMGIKTAIVGAVYRKALVTNSARKSTV 420
Db 361 WQGFYTVLAEVTAACLOTIVLHOYFHI CFVSGMGIKTAIVGAVYRKALVTNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATVYINMTSAPLQVITLALYLLMLLGSVLAGVAVMYLMEVFN 480
Db 421 GEIVNLSVDAQRFMDLATVYINMTSAPLQVITLALYLLMLLGSVLAGVAVMYLMEVFN 480
QY 481 AVMAKTKTYQVAMKSKDNRIKIMNEILNGIKVLYAMELAFKDYALAIROBELKYLK 540
Db 481 AVMAKTKTYQVAMKSKDNRIKIMNEILNGIKVLYAMELAFKDYALAIROBELKYLK 540
QY 541 KSAIYLSAVGTFWTCTPFLVALCTFAVYVITDENNIIDAQAFPSLALFNILRFPNITLP 600
Db 541 KSAIYLSAVGTFWTCTPFLVALCTFAVYVITDENNIIDAQAFPSLALFNILRFPNITLP 600
QY 601 MVISIYQASVSLRLNIFLSHELEPDSIERRPVKQCGGNSITVRNATFTMARSDPT 660
Db 601 MVISIYQASVSLRLNIFLSHELEPDSIERRPVKQCGGNSITVRNATFTMARSDPT 660
QY 661 LNGITFSIPREGALVAVVQVCGKSSLSALLAEMDVYEGHVAIKGSVAVYVPOQAMIQND 720
Db 661 LNGITFSIPREGALVAVVQVCGKSSLSALLAEMDVYEGHVAIKGSVAVYVPOQAMIQND 720
QY 721 SLRENILFGCQLEPPYRSVYQAQALLPDLEILPSSGRTEIGEGKNVLSGGQKRVSLAR 780
Db 721 SLRENILFGCQLEPPYRSVYQAQALLPDLEILPSSGRTEIGEGKNVLSGGQKRVSLAR 780
QY 781 AVYSNADITLFPDDPLASVADAVGHIFENYVIGPGMLKNKTRILVTHSMSTLPQVDYIIV 840
Db 781 AVYSNADITLFPDDPLASVADAVGHIFENYVIGPGMLKNKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGGKISEMSGYOELIARDAFAEFILRTYASTEOBDAEENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGGKISEMSGYOELIARDAFAEFILRTYASTEOBDAEENGVTGVSQPGKEAKOMENGM 900
QY 901 LVYDSAGKOLQROLSSSSSYSGDISRHNSYTAELQAKAEKEETWKLMEADKATQGVYKL 960
Db 901 LVYDSAGKOLQROLSSSSSYSGDISRHNSYTAELQAKAEKEETWKLMEADKATQGVYKL 960

QY		961	SVYVMDYKMAIGEIFEFLSLFLEMCNCHVASLANSWYLSMTDDEPIYNGOHEHTKRLSYSG	1020
Dd		961	SVYVMDYKMAIGEIFEFLSLFLEPICNHVAALSNWYLSMTDDPIYNGOHEHTKRLSYSG	1020
QY		1021	ALGISOGIAVFQYSNAVSIGGILASRCIHAVDLHSILSPMSFFERTPSGLVNRFSKEL	1080
Dd		1021	ALGISOGIAVFQYSNAVSIGGILASRYLHVLDLHSILSPMSFFERTPSGLVNRFSKEL	1080
QY		1081	DTVDSDMIPRVIYKMFPGSLFNVTGACIVILLATPILAIITIIPLGLIYFFVGQRPVYASSROL	1140
Dd		1081	DTVDSDMIEPVIMFGSFLFNVTGACIVILLATPILAIITIIPLGLIYFFVGQFPVYASSROL	1140
QY		1141	KRLSEVSRSPPVSHNETLLGVSVIRAFEEORERFHOSDKVDENOKRAYESIYANRWLA	1200
Dd		1141	KRLSEVSRSPPVSHNETLLGVSVIRAFEEDERFHOSSDKVDENOKRAYESIVANRWLA	1200
QY		1201	VRLFCVGNCLVLPALPAVISRHSHLSAGLVGLSYSLSQVTTYLNMLVRMSEMETINIVA	1260
Dd		1201	VRLFCVGNCLVLPALPAVISRHSHLSAGLVGLSYSLSQVTTYLNMLVRMSEMETINIVA	1260
QY		1261	VERLKEYSSTEKAEAWQIQETAPRSSPOVGRVEFRNYCLARYBDLDFVLRIHINTINGG	1320
Dd		1261	VERLKEYSSTEKAEAWQIOETAPRNMPQVGRVEFRNYCLAREBDDLDFVLRIHINTINGG	1320
QY		1321	EKVYGVGTGAGKSLLTGLFPIINSAGEITIIIGGINAKIGIHLDRKTIITIPODPVL	1380
Dd		1321	EKVYGVGTGAGKSLLTGLFPIINSAGEITIIDGINARIGIHLDRKTIITIPODPVL	1380
QY		1381	SGSLRMNLDPFSQYSDEEWWTSLSLAHLKDFVSALPDKLDEHCAEGENLSCGOROLVCL	1440
Dd		1381	SGSLRMNLDPFSQYSDEEWWTSLBLAHLKGFPSALPDGLDEHCAEGENLSCGOROLVCL	1440
QY		1441	ARALLRKTKILVLDATAAVDLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVYL	1500
Dd		1441	ARALLRKTKIIVLDATAAVDLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVYL	1500
QY		1501	DKGEIOEQGAPSDLLQQRGLFPMAXKDAGLV	1531
Dd		1501	DKGEIOEQGAPSDLLQQRGLEPYNMARDAGLV	1531
RESULT 5				
Q9UO97_HUMAN				
ID	Q9UO97_HUMAN	PRELIMINARY;	PRT;	1459 AA.
AC	Q9UO97_			
DT	01-MAY-2000	(TrEMBLrel. 13, Created)		
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		
DE	Multidrug resistance protein (Fragment).			
GN	Name=MRP;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;			
OC	Homo.			
NCBI_TaxID=9606;	[1]			
NUCLEOTIDE SEQUENCE.				
RX	MEDLINE=96008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;			
RA	Grant C.E., Kurz E.V., Cole S.P.C., Deeley R.G.;			
RT	"Analysis of the intron-exon organization of the human multidrug-			
RT	resistance protein gene (MRP) and alternative splicing of its mRNA.";			
RL	Genomics 45:368-378(1997).			
CC	-1- SIMILARITY: Belongs to the ABC transporter family.			
DR	EMBL; AF022853; AAB83980.1; -; Genomic DNA.			
DR	EMBL; AF022824; AAB83980.1; JOINED; Genomic DNA.			
DR	EMBL; AF022825; AAB83980.1; JOINED; Genomic DNA.			
DR	EMBL; AF022826; AAB83980.1; JOINED; Genomic DNA.			
DR	EMBL; AF022828; AAB83980.1; JOINED; Genomic DNA.			
DR	EMBL; AF022830; AAB83980.1; JOINED; Genomic DNA.			
DR	EMBL; AF022833; AAB83980.1; JOINED; Genomic DNA.			
DR	EMBL; AF022834; AAB83980.1; JOINED; Genomic DNA.			
DR	EMBL; AF022836; AAB83980.1; JOINED; Genomic DNA.			

	DR	EMBL; AF022838; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022846; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022847; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022845; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022844; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022843; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022842; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022841; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022840; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022839; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022838; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022837; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022836; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022835; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022834; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022833; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022832; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022831; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022830; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022829; AAB83980.1; JOINED; Genomic DNA.
	DR	EMBL; AF022828; AAB83980.1; JOINED; Genomic DNA.
	DR	HSSP; P08716; IMT0.
	DR	Ensembl; ENSG00000103222; Homo sapiens.
	DR	GO; GO:0016021; C:integral to membrane; IEA.
	DR	GO; GO:0005524; F:ATP binding; IEA.
	DR	GO; GO:0016887; F:ATPase activity; IEA.
	DR	GO; GO:0042628; F:ATPase activity, coupled to transmembrane m. . ; IEA.
	DR	GO; GO:0000166; F:nucleotide binding; IEA.
	DR	GO; GO:0006810; P:transport; IEA.
	DR	InterPro; IPRO03593; AAA ATPase.
	DR	InterPro; IPRO11527; ABC_membrane_1.
	DR	InterPro; IPRO01140; ABC_TM_transpt.
	DR	InterPro; IPRO03439; ABC_cranp_like.
	DR	InterPro; IPRO05292; MRP_assoc.
	DR	InterPro; IPRO00719; Prot_kinase.
	DR	Pfam; PF00664; ABC_membrane; 2.
	DR	Pfam; PF00005; ABC_cran; 1.
	DR	ProDom; PD000006; ABC_transporter; 1.
	DR	SMART; SM00382; AAA; 2.
	DR	TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
	DR	PROSITE; PS02929; ABC_TMIF; 2.
	DR	PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
	DR	PROSITE; PS08993; ABC_TRANSPORTER_2; 2.
	DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
	KW	ATP-binding; Nucleotide-binding.
	FT	NON TER
	SQ	SEQUENCE 1459 AA; 163831 MW; 6A016A5A6E7AA95C CRC64;
	Query Match	74.3%; Score 7444; DB 2; Length 1459;
	Best Local Similarity	96.3%; Pred. No. 0;
	Matches 1459; Conservative % 0; Mismatches 0; Indels 56; Gaps 1;	
Oy	17	DMMVTNNTSNDPFKFCFQNTVTVVWPCEFYLMACFPYFLYSRRDRTGIOMTPINKTKTA 76
Db	1	DMMVTNNTSNDPFKFCFQNTVTVVWPCEFYLMACPFFYLISRDRDTGIOMTPINKTKTA 60
Oy	77	LGFLLIMIVCMADLYFSFWERSRGIFLAPVLAVSTLTGLITTLATPLTOLERRKGVOSSG 136
Db	61	LGFLLIMIVCMADLYFSFWERSRGIFLAPVLAVSTLTGLITTLATPLTOLERRKGVOSSG 120
Oy	137	IMLTFFWLVALVCALAIIRSKIKMTALKEDAQVDLEFRDITFYVYFSLLIQTVLSCFSRSP 196
Db	121	IMLTFFWLVALVCALAIIRSKIKMTALKEDAQVDLEFRDITFYVYFSLLIQTVLSCFSRSP 180
Oy	197	LFSETHIDPNPCPPSSASFLSRITFWMITGTIYNGYROPLEGSDLSLNKEDTSBOVVVP 256
Db	181	LFSETHIDPNPCPPSSASFLSRITFWMITGTIYNGYROPLEGSDLSLNKEDTSBOVVVP 240
Oy	257	LVKMVKKECAKTRQPKVYVVSXKDPAQPKESSTVDANEEVEALIVSPKENVPSLFFKV 316
Db	241	LVKMVKKECAKTRQPKVYVVSXKDPAQPKESSTVDANEVEALIVSPKENVPSLFFKV 300
Oy	317	LYKTGGYFLMSFFPKAIHDLMFSGQILKLTIKYNDTKRADPDGYFTYVILLVFYACL 376

Db 301 LYKTFGEYFLMSFFPKAIHIDLMFSGPOLIKLIKFNVDTPADPMQGYFTYVTLFTVACL 360
Qy 377 QTVLHVOYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTVGEIVNLMSVDAORFMD 436
Db 361 QTVLHVOYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTVGEIVNLMSVDAORFMD 420
Qy 437 LATYNNIWSAPLOVILALYILMLNLGPSVLAGVAVMLAMPVNAVAMAKTKTYQVAHMK 496
Db 421 LATYNNIWSAPLOVILALYILMLNLGPSVLAGVAVMLAMPVNAVAMAKTKTYQVAHMK 480
Qy 497 SKDRIKLMBEILNGIKVLKLYAMELIAFKDVKLAIROBELKVLKKSATLSVGFPTWCT 556
Db 481 SKDRIKLMBEILNGIKVLKLYAMELIAFKDVKLAIROBELKVLKKSATLSVGFPTWCT 540
Qy 557 PFLVALCTFAVYVITIDENNIIIDAOTAFVSLALFNILRPPLNILPMVSIISVQASVSLKRL 616
Db 541 PFLVALCTFAVYVITIDENNIIIDAOTAFVSLALFNILRPPLNILPMVSIISVQASVSLKRL 600
Qy 617 RIFLSHELEBDSIERRPVKDGGTNSITVRNATFTMARSDPPLNGITFSIPGALVAV 676
Db 601 RIFLSHELEBDSIERRPVKDGGTNSITVRNATFTMARSDPPLNGITFSIPGALVAV 660
Qy 677 VGOVGGCKSSLALMLBMDKVBGHVATKGSVAVVPQAMIONSLARENILFGQLEBPY 736
Db 661 VGOVGGCKSSLALMLBMDKVBGHVATKGSVAVVPQAMIONSLARENILFGQLEBPY 720
Qy 737 YRSYIOACALPDLIELPSGDRTEIGEKVNLSSGGQKORVSLARAVSNADIVLFDPELS 796
Db 721 YRSYIOACALPDLIELPSGDRTEIGEKVNLSSGGQKORVSLARAVSNADIVLFDPELS 747
Qy 797 AVDAHVGKHLFENVIGPKMLKNKTRILIVTHSMSTLPQVDVIVMSSGKISEMSYOEHL 856
Db 748 AVDAHVGKHLFENVIGPKMLKNKTRILIVTHSMSTLPQVDVIVMSSGKISEMSYOEHL 784
Qy 857 ARDGAFAEPLRTYASTROBODAEENGVTGSGPGGEAKOMNGMLYTDSACKQIOROLSS 916
Db 785 ARDGAFAEPLRTYASTROBODAEENGVTGSGPGGEAKOMNGMLYTDSACKQIOROLSS 844
Qy 917 SSSYSGDISRRHNSTAELOKAEAKKEETWKMEADKAOTGOVKLSVYWDYMKAIGLFTSF 976
Db 845 SSSYSGDISRRHNSTAELOKAEAKKEETWKMEADKAOTGOVKLSVYWDYMKAIGLFTSF 904
Qy 977 LSLIFLPMCNHVSALASNWTLSLMTDDPIVNGSTOEHTKRLSVYGAIGISOGIAVFGYSMA 1036
Db 905 LSLIFLPMCNHVSALASNWTLSLMTDDPIVNGSTOEHTKRLSVYGAIGISOGIAVFGYSMA 964
Qy 1037 VSIIGIILASRCLHVDLHSLIRSPSPFERTPSGNLVNRFSEKELDTVDSMTPEVYKMFEG 1096
Db 965 VSIIGIILASRCLHVDLHSLIRSPSPFERTPSGNLVNRFSEKELDTVDSMTPEVYKMFEG 1024
Qy 1097 SLFNVIGACIYILATPIAIIIPPLGIYFFVQRFYVASSROQLKRLSVSRSPYSHFN 1156
Db 1025 SLFNVIGACIYILATPIAIIIPPLGIYFFVQRFYVASSROQLKRLSVSRSPYSHFN 1084
Qy 1157 ETLILGVSIVIRAFEEQERFIHOSDLKVDENQKAYYPSIVANRWLAVRLCEVCAGCIYLFAL 1216
Db 1085 ETLILGVSIVIRAFEEQERFIHOSDLKVDENQKAYYPSIVANRWLAVRLCEVCAGCIYLFAL 1144
Qy 1217 FAVISRSLSLAGLVLSYSLOTTYLANMLVRNMSSEMETNIVAVERLKEYSETEKEAPW 1276
Db 1145 FAVISRSLSLAGLVLSYSLOTTYLANMLVRNMSSEMETNIVAVERLKEYSETEKEAPW 1204
Qy 1277 QIOETAPSSWPQYGRVFRFRNYCLRYREDLDFVLRHINVTIINGEKVGIORGTAGKSSSL 1336
Db 1205 QIOETAPSSWPQYGRVFRFRNYCLRYREDLDFVLRHINVTIINGEKVGIORGTAGKSSSL 1264
Qy 1337 TLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLPFSGSLRMLDPPSQYSD 1396
Db 1265 TLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLPFSGSLRMLDPPSQYSD 1324
Qy 1397 EEVWTSJLELAHKDFVSLPDKLDHECBGGENLSVGOQOLVCLARALLRTKTIIVLDEA 1456
Db 1325 EEVWTSJLELAHKDFVSLPDKLDHECBGGENLSVGOQOLVCLARALLRTKTIIVLDEA 1384

Qy 1457 TAAVDLETDDLIOSTIRFOEDCTVLTIAHRLNTIMDYTVIYVDKGEIOEGAPSDLLQ 1516
Db 1385 TAAVDLETDDLIOSTIRFOEDCTVLTIAHRLNTIMDYTVIYVDKGEIOEGAPSDLLQ 1444
Qy 1517 ORGLFYSMARDAGLV 1531
Db 1445 ORGLFYSMARDAGLV 1459

RESULT 6
Q9UQAO_HUMAN PRELIMINARY, PRT, 1456 AA.
ID Q9UQAO_HUMAN PRELIMINARY, PRT, 1456 AA.
AC Q9UQAO;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2004 (Tremblrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
RT resistance protein gene (MRP) and alternative splicing of its mRNA.";
RT Genomics 45:368-378(1997).
CC -I- SIMILARITY: Belongs to the ABC transporter family.
CC -I- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AF022825; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022824; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022826; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022827; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022829; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022831; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022832; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022833; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022835; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022837; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022846; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022847; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022848; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022849; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022845; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022844; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022843; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022842; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022841; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022840; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022838; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022852; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022851; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022850; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022849; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022848; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022836; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022834; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022832; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022830; AAB83981.1; JOINED; Genomic DNA.
DR EMBL: AF022828; AAB83981.1; JOINED; Genomic DNA.
HSSP: P08716, 1MT0.
DR Ensemble: ENSG00000103222; Homo sapiens.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005524; F: ATP binding; IEA.
DR GO: GO:0016887; F: ATPase activity; IEA.
DR GO: GO:0042626; F: ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000166; F: nucleotide binding; IEA.
DR GO: GO:0006810; P: transport; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR011527; ABC_membrane_1.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.

DR InterPro: IPR005292; MRP_assoc.
 DR InterPro: IPR000719; Prot_kinase.
 DR Pfam: PF00664; ABC_membrane_2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PD000006; ABC_transporter; 2.
 DR SMART: SM00382; AAA; 2.
 DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE: PS00929; ABC_TMIF; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR ATP-binding; Nucleotide-binding.
 FT NON_TER 1
 SQ SEQUENCE 1456 AA; 163233 MW; 8D89AAB2BC481F2 CRC64;

Query Match 74.1%; Score 7419.5; DB 2; Length 1456;
 Best Local Similarity 96.1%; Pred. No. 0;
 Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;

QY 17 DMNVNTSNPDPFKCFONTVLVWPCFYLMACFPFYLYSRHGRYIQMTPIKTKTA 76
 DB 1 DMNVNTSNPDPFKCFONTVLVWPCFYLMACFPFYLYSRHGRYIQMTPIKTKTA 60

QY 77 LGFLMIVCWADLFYSFMERSRGIFLAPVFLVSPTLIGITTLATFLIOLERRKGVSSG 136
 DB 61 LGFLMIVCWADLFYSFMERSRGIFLAPVFLVSPTLIGITTLATFLIOLERRKGVSSG 120

QY 137 IMTLFWLVALCALAIIRSKIMTALKEDAQVDLFRDITFYVYFSLIIQLVSCFSRSP 196
 DB 121 IMTLFWLVALCALAIIRSKIMTALKEDAQVDLFRDITFYVYFSLIIQLVSCFSRSP 180

QY 197 LFBETIHDPNCPRESSASPLSRITFWMTITGLIVGYPQLPGLSLMNLNEDTSEQVVPV 256
 DB 181 LFBETIHDPNCPRESSASPLSRITFWMTITGLIVGYPQLPGLSLMNLNEDTSEQVVPV 240

QY 257 LVNWKKECAKTRKQPKVYVSSKDPQPKSSKVDNNEVEALIVSPKEMNPSLFKV 316
 DB 241 LVNWKKECAKTRKQPKVYVSSKDPQPKSSKVDNNEVEALIVSPKEMNPSLFKV 300

QY 317 LYTFPGFYFLMSFFPKAIDHLMFSGPQILKLIFVNDTKAPDMQGYFTVLLFTVACL 376
 DB 301 LYTFPGFYFLMSFFPKAIDHLMFSGPQILKLIFVNDTKAPDMQGYFTVLLFTVACL 360

QY 377 QTVLHYEHCIFSGGRITKAVIGAYRKALVTNARKSSTYGEIVNLMSVDAQFMD 436
 DB 361 QTVLHYEHCIFSGGRITKAVIGAYRKALVTNARKSSTYGEIVNLMSVDAQFMD 420

QY 437 LATYINNIWSPAPLOVIALYILMLNLGSPVLAGAVAVWLVAVNAVMAKTKTYQVAHMK 496
 DB 421 LATYINNIWSPAPLOVIALYILMLNLGSPVLAGAVAVWLVAVNAVMAKTKTYQVAHMK 480

QY 497 SKDNRIKLMEIINGIKVLKLYAMELAFKDKVLAIRQEBLKKSAVLSAVGTFTVCT 556
 DB 481 SKDNRIKLMEIINGIKVLKLYAMELAFKDKVLAIRQEBLKKSAVLSAVGTFTVCT 540

QY 557 PELVALCTPVTYVTTIDENNLIDQATAVSLAFNIIILFPNIIIPMTVSSIVQASVSKRL 616
 DB 541 PELVALCTPVTYVTTIDENNLIDQATAVSLAFNIIILFPNIIIPMTVSSIVQASVSKRL 600

QY 617 RIFLSHEELPDSIERPVDGGGNSITVRNATFTARSDDPLNLTISIEBGAIVAV 676
 DB 601 RIFLSHEELPDSIERPVDGGGNSITVRNATFTARSDDPLNLTISIEBGAIVAV 660

QY 677 VGVGCGKSSLLSALLAEMDKVEGHVAKIGSVAVVPOQAMIQNDSLRENTILFGQOLEEPP 736
 DB 661 VGVGCGKSSLLSALLAEMDKVEGHVAKIGSVAVVPOQAMIQNDSLRENTILFGQOLEEPP 688

QY 737 YREVIQACALLPDLLEILPSGDRTEIGRKVNLSSGQKQVSLARAVYSNADIYLFDPPLS 796
 DB 689 YREVIQACALLPDLLEILPSGDRTEIGRKVNLSSGQKQVSLARAVYSNADIYLFDPPLS 721

QY 797 AYDAHVAKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVVIIVMSGKISSEMGSYOELL 856
 DB 797 AYDAHVAKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVVIIVMSGKISSEMGSYOELL 856

DB 722 AYDAHVAKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVVIIVMSGKISSEMGSYOELL 781

QY 857 ARDGAFAEPLRTYASTBOBDAEENGTVGVSGPCKEKKOMENGLVYDSAGKOLQROLS 916
 DB 782 ARDGAFAEPLRTYASTBOBDAEENGTVGVSGPCKEKKOMENGLVYDSAGKOLQROLS 841

QY 917 SSSYSGDISHHNSTAELOKAEAKKEETWLMKENDKQOTQVXLVYVMDYKAIGLFTSF 976
 DB 842 SSSYSGDISHHNSTAELOKAEAKKEETWLMKENDKQOTQVXLVYVMDYKAIGLFTSF 901

QY 977 LSIFLEMCNVSALASNYMLSLMTDDEIVNGTOEHTKRLSVYGALGISGIAVFGYSMA 1036
 DB 902 LSIFLEMCNVSALASNYMLSLMTDDEIVNGTOEHTKRLSVYGALGISGIAVFGYSMA 961

QY 1037 VSIIGILASRCLHYDLHSLIRSPMSFEETPSGNLVNRSKELDTYDSMIPEYIKMPMG 1096
 DB 962 VSIIGILASRCLHYDLHSLIRSPMSFEETPSGNLVNRSKELDTYDSMIPEYIKMPMG 1021

QY 1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFPQRFVAVASSROLKLESVSRSPVSHFN 1156
 DB 1022 SLFNVIGACIVILLATPIAIIIPPLGLIYFPQRFVAVASSROLKLESVSRSPVSHFN 1081

QY 1157 ETLIGSVIRAFBOERFIHQSDLKVDENOKAYYPSIVANRWLAVRLECVGNCTVLPAL 1216
 DB 1082 ETLIGSVIRAFBOERFIHQSDLKVDENOKAYYPSIVANRWLAVRLECVGNCTVLPAL 1141

QY 1217 FAVISRHSLSAGLVGSVSLQVTTYLNLVMSSEMETNIVAVERLKEYSSETEKAPW 1276
 DB 1142 FAVISRHSLSAGLVGSVSLQVTTYLNLVMSSEMETNIVAVERLKEYSSETEKAPW 1201

QY 1277 QIOTAPSSWPQVGRPEFPNVCARREDLPFLRHNTVINGEKVGIYRGAGKSSL 1336
 DB 1202 QIOTAPSSWPQVGRPEFPNVCARREDLPFLRHNTVINGEKVGIYRGAGKSSL 1261

QY 1337 TLGLFRINSEAGEIIIDGINIAKIGHDRFKITIIIPQPVLFSGSLRANLDPFSQYSD 1396
 DB 1262 TLGLFRINSEAGEIIIDGINIAKIGHDRFKITIIIPQPVLFSGSLRANLDPFSQYSD 1321

QY 1397 EEWVTSLELAHLDQFVSALPDKLDHCEGEGNLSVGOQOLVCLARALLKTKILVLEDA 1456
 DB 1322 EEWVTSLELAHLDQFVSALPDKLDHCEGEGNLSVGOQOLVCLARALLKTKILVLEDA 1381

QY 1457 TAAVDETDLIOSTRTQPEDCTVLTARLNTIDYFTVYLDEGEIOEGAPSDLQ 1516
 DB 1382 TAAVDETDLIOSTRTQPEDCTVLTARLNTIDYFTVYLDEGEIOEGAPSDLQ 1441

QY 1517 QRGLEYSMAKAGLV 1531
 DB 1442 QRGLEYSMAKAGLV 1456

RESULT 7
 Q6UR05 CANPA
 ID Q6UR05 CANPA PRELIMINARY; PRT; 1531 AA.
 AC Q6UR05;
 DT 05-JUL-2004 (Tremblrel. 27, Created)
 DT 05-JUL-2004 (Tremblrel. 27, last sequence update)
 DT 05-JUL-2004 (Tremblrel. 27, last annotation update)
 DE Multidrug resistance-associated protein 1.
 GN Name=MRP1;
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22404678; PubMed=12516967;
 RA Ma L., Pratt S.B., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;
 RT "Identification and characterization of the canine multidrug
 resistance-associated protein.";
 RL Mol. Cancer Ther. 1:1335-1342(2002).
 [2]

RP NUCLEOTIDE SEQUENCE.
Ma L., Prate S.E., Cao J., Danczig A.H., Moore R.E., Slapak C.A.;
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SMILARITY: Belongs to the ABC transporter family.
DR EMBL: AY631728; AAO3148.1; -; mRNA.
DR EMBL: ENSCAFG00000018208; Cantis familiaris.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042666; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR011527; ABC membrane 1.
DR InterPro: IPR01140; ABC TM transp.
DR InterPro: IPR03439; ABC transp_like.
DR InterPro: IPR005292; ABC assoc.
DR InterPro: IPR000719; Prot_kinase.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PSS0929; ABC_TMtr; 2.
DR PROSITE: PSS0211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PSS0893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PSS0107; PROTEIN KINASE_ATP; UNKNOWN 1.
KW ATP-binding; Nucleotide-binding; Repeat; Transport.
SQ SEQUENCE 1531 AA; 171792 MW; 746361A71C6158BD CRC64;

Query Match 72.6%; Score 7272; DB 2; Length 1531;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;

QY 1 MALRGPSADSDSLMDMNTWNTSNPDFTKQONTVVMVPCFYLACPFYFLYSRH 60
DB 1 MALRGPSADSDSLMDMNTWNTSNPDFTKQONTVVMVPCFYLACPFYFLYSRH 60
QY 61 DRGIQMTPLNKTALGFLIMVCMADLFYSFMERSGIFLAPVFLVSPILGITLLA 120
DB 61 DRGIQMTPLNKTALGFLIMVCMADLFYSFMERSGIFLAPVFLVSPILGITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTWLVALCALILNSKIMTALKEDAQVDLPDITPYVFS 180
DB 121 TFLIQLERRKGVSSGIMLTWLVALCALILNSKIMTALKEDAQVDLPDITPYVFS 180
QY 181 LLLIQLVLSGSDSPLESETIHDNPNCPSSASFLSITFMWITGLIVRGRQPLESD 240
DB 181 LLLIQLVLSGSDSPLESETIHDNPNCPSSASFLSITFMWITGLIVRGRQPLESD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKECAKTRKQPVKVVS KDPAQKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKECAKTRKQPVKVVS KDPAQKSSKYDANEVEAL 300
QY 301 IVKSPQKEWNSLTKVLYKTGPFYFLMSFFPKAIDHLMFSGPOILKLLIKFVNDKAP 360
DB 301 IVKSPQKEWNSLTKVLYKTGPFYFLMSFFPKAIDHLMFSGPOILKLLIKFVNDKAP 360
QY 361 WQGFYVTVLFTVACLOTLVHOYFHICFVSGMRKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYVTVLFTVACLOTLVHOYFHICFVSGMRKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLSVDQRFMDLATYINMIWSAPLOVILALYLMLNGPSVLAVVMVLMVFN 480
DB 421 GEIYNLSVDQRFMDLATYINMIWSAPLOVILALYLMLNGPSVLAVVMVLMVFN 480
QY 481 AVMAKTKTYGVANMKSNDRIKLMNEILNGIKVILKYAMEIAFKDXYLAIROBELKYLK 540
DB 481 AVMAKTKTYGVANMKSNDRIKLMNEILNGIKVILKYAMEIAFKDXYLAIROBELKYLK 540
QY 541 KSAATLSAVGTFTWCTPFLVALCTPFAVVTIDENNIIDNAQAFVSLAFNLRPFIATLP 600
DB 541 KSAATLSAVGTFTWCTPFLVALCTPFAVVTIDENNIIDNAQAFVSLAFNLRPFIATLP 600

QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVXOGGNTSNVTNNATFTWARSDEPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVXOGGNTSNVTNNATFTWARSDEPT 660
QY 661 LNGITFSPGALVAVVGVQCGKSSLSLALNEMDVBEHVAIKGSVAVYPOQAWIQND 720
DB 661 LSGITFSPGSLVAVVGVQCGKSSLSLALNEMDVBEHVAIKGSVAVYPOQAWIQND 720
QY 721 SLRENIIFGQLEBPYRSYIOACALLPDEIIPSGDRTEIGEGKVALSGGQKORVSLAR 780
DB 721 SLRENIIFGQLEBPYRSYIOACALLPDEIIPSGDRTEIGEGKVALSGGQKORVSLAR 780
QY 781 AVYSNADIIYLFDDPLSAVDHVGKHPENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYCDSDIYLFDDPLSAVDHVGKHPENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGKISSEMSYQELRLARDGAFAPFLTASTBOEQAENGVTGVSGPKKAKOMENGM 900
DB 841 MTGKISSEMSYQELRLARDGAFAPFLTASTBOEQAENGVTGVSGPKKAKOMENGM 900
QY 901 LVTDASGKOLOROLSSSSYSGDISRHNSTABLOKAEAKKEETWKLMEADKAOTGVKL 960
DB 901 LVTDVAGKOLOROLSSSSYSGDYSRHTSTABLOKAGPKNEADMKLVEADKAOTGVKL 960
QY 961 SVYWDYKKAIGLFIPLSIFLPMCNSVALASNYMGLMTDDPIVNGTOEHTKVLSSVG 1020
DB 961 SVYWDYKKAIGLFIPLSIFLPMCNSVALASNYMGLMTDDPIVNGTOEHTKVLSSVG 1020
QY 1021 ALGISQIAVFGYSMAVSTIGIILASRCLHVDLHSILRSMPFPERPSSGLVNRFEKL 1080
DB 1021 ALGISQIAVFGYSMAVSTIGIILASRCLHVDLHSILRSMPFPERPSSGLVNRFEKL 1080
QY 1081 DTVDMSIPEYIKMPGSLFNVIAGCIYLLATPIAIIIPPLGIYFFVQFYAASSROL 1140
DB 1081 DTVDMSIPEYIKMPGSLFNVIAGCIYLLATPIAIIIPPLGIYFFVQFYAASSROL 1140
QY 1141 KRLESVSRSPYSHFNELLGVSVIRAFBEQERFIHQSDLKVDENQKAYYSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNELLGVSVIRAFBEQERFIHQSDLKVDENQKAYYSIVANRWLA 1200
QY 1201 VRELCVNCIYLPALRAVISRHSLSAGLVLSVSYLOVTVLNMVRSSEMETNIVA 1260
DB 1201 VRELCVNCIYLPALRAVISRHSLSAGLVLSVSYLOVTVLNMVRSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOEMAPPSWPOGRVFRYCYLRREDLPVLRHINTNGG 1320
DB 1261 VERLKEYSETEKEAPMOIOEMAPPSWPOGRVFRYCYLRREDLPVLRHINTNGG 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVL 1380
DB 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVL 1380
QY 1381 SGIIRANLDPFSQYSDDEWTSLELAHKDFVSAAPKLDHECAEGENISVGOEOLVCL 1440
DB 1381 SGIIRANLDPFSQYSDDEWTSLELAHKDFVSAAPKLDHECAEGENISVGOEOLVCL 1440
QY 1441 ARAALRRTKILIVDEAATAVDLETDLLQSTIRQFDDCTVLTATRLNTIMDTRIYVL 1500
DB 1441 ARAALRRTKILIVDEAATAVDLETDLLQSTIRQFDDCTVLTATRLNTIMDTRIYVL 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIRREGQPSDILQORGLFYSMAKDAGLV 1531

RESULT 8
059G19 HUMAN
ID 059G19 HUMAN PRELIMINARY; PRT; 1439 AA.
AC 059G19
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)

DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Multidrug resistance protein 1.
GN Name=MRP1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;
RA Teguchi Y., Saeki K., Komano T.;
RT "Functional analysis of MRP1 cloned from bovine."
RL FBS Lett. 521:211-213(2002).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AB082124; BAC15550.1; -; mRNA.
DR HSSP; P08716; 1MT0.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC TM transp.
DR InterPro; IPR003439; ABC transp_like.
DR Pfam; PF00064; ABC_membrane; 2.
DR Pfam; PF00064; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRfams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS0929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1530 AA; 171666 MW; AAE4F92ED7832703 CRC64;

Query Match 72.2%; Score 7230.5; DB 2; Length 1530;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;

DY 1 MALSGFSGADGSDPLMDNMNTWNTSNPDFTKCFONTVLVWVPCFTLACFPFYLYSRH 60
1 MALDPCFSDGSDLFWEWNVNTWNTSNPDFTKCFONTVLVWVPCFTLWVCFPFYLYSSH 60
Dy 61 DRGYIOMTPLAKTKTALGFLIMIVCMADLFYSFMRSRGIFLAPVFLVSPILLGITTLA 120
Dy 61 DRGYIOMTHLAKATLALGFLIMIVCMADLFYSFMRSRGKGLAPVFLVSPILLGITTLA 120
Dy 121 TFLIQLERRKGVSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLPDITFYVRS 180
Dy 121 TFLIQLERRRQVSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLPDITFYVRS 180
Dy 121 TFLIQLERRRQVSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLPDITFYVRS 180
Dy 181 LLLIQLVLSGSDSPLESETIHDPNCPSSASFLSRTFMWITGLVGRVGRPLEGSD 240
Dy 181 LLLIQLVLSGSDSPLESETIHDPNCPSSASFLSRTFMWITGLVGRVGRPLEGSD 240
Dy 181 LLLIQLVLSGSDSPLESETIHDPNCPSSASFLSRTFMWITGLVGRVGRPLEGSD 240
Dy 241 LMSLNKEDTSQVVPVLVKNKKKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
Dy 241 LMSLNKEDTSQVVPVLVKNKKKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
Dy 301 IVKSPQKEMNSLKFVLYKTFGYPYLMSPFKALHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Dy 301 IVKSPQKEMNSLKFVLYKTFGYPYLMSPFKALHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Dy 301 IVKSPQKEMNSLKFVLYKTFGYPYLMSPFKALHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Dy 361 WQGFYFTLLFVTAACLOTLLVHGYPHICFVSGMRKTAIVIGAVYKALVITNSARKSTV 420
Dy 361 WQGFYFTLLFVTAACLOTLLVHGYPHICFVSGMRKTAIVIGAVYKALVITNSARKSTV 420
Dy 361 WQGFYFTLLFVTAACLOTLLVHGYPHICFVSGMRKTAIVIGAVYKALVITNSARKSTV 420

Dy 421 GEIVNLSVDAQRFMDIATYINNIWSAPLOVIALYLLMLNGLPSVLAVAVMLANPVN 480
Dy 421 GEIVNLSVDAQRFMDIATYINNIWSAPLOVIALYLLMLNGLPSVLAVAVMLANPVN 480
Dy 481 ANWAKTKTYOVAMHSKDNRIKLMELLNGIVLKLAYELKPKQVLAIRQELVVK 540
Dy 481 ANWAKTKTYOVAMHSKDNRIKLMELLNGIVLKLAYELKPKQVLAIRQELVVK 540
Dy 541 KSAVLSAVGTFTWCFTFVALCTFAVYVITDENNIIDAQTAFLALFNILRFNLIILP 600
Dy 541 KSAVLSAVGTFTWCFTFVALCTFAVYVITDENNIIDAQTAFLALFNILRFNLIILP 600
Dy 541 KSAVLSAVGTFTWCFTFVALCTFAVYVITDENNIIDAQTAFLALFNILRFNLIILP 600
Dy 601 MVISSIVQASVSLKRLRIPLSHEELEPDSIERRPVKDGCGTNSITVNAFTWASDPPT 660
Dy 601 MVISSIVQASVSLKRLRIPLSHEELEPDSIERRPVKDGCGTNSITVNAFTWASDPPT 660
Dy 661 LMGITFSEPGALVAAVGVGCGKSSILSLALAEMLDVBEGHVAIKGSVAVPOQAMTQND 720
Dy 661 LMGITFSEPGALVAAVGVGCGKSSILSLALAEMLDVBEGHVAIKGSVAVPOQAMTQND 720
Dy 721 SLRENILFGQLEBPYRSVYIQAALLPDLIELPSGDRTEIGEGVNLASGQKORVSLAR 780
Dy 721 SLRENILFGQLEBPYRSVYIQAALLPDLIELPSGDRTEIGEGVNLASGQKORVSLAR 780
Dy 781 AVYSNADYLPDPLSAVDHVGKHIFENYIGPKMLAKNTRILVTHSMGTYLPQVDYIIV 840
Dy 781 AVYCDSDVYLLDPLSAVDHVGKHIFENYIGPKMLAKNTRILVTHSMGTYLPQVDYIIV 840
Dy 841 MSGGKISEMSYQELLARDGAPAEFLRTVASTQEQDAAENGVGVSGPGKAAKQMGNGM 900
Dy 841 MSGGKISEMSYQELLARDGAPAEFLRTVASTQEQDAAENGVGVSGPGKAAKQMGNGM 900
Dy 901 LVTSACKQOROLSSSSSYSGDISRRHNSTAELOKAEAKKEETKMLEADKATQGVKL 960
Dy 901 LVTSACKQOROLSSSSSYSGDISRRHNSTAELOKAEAKKEETKMLEADKATQGVKL 960
Dy 901 LVTSACKQOROLSSSSSYSGDISRRHNSTAELOKAEAKKEETKMLEADKATQGVKL 960
Dy 961 SVYDYNKAIQGLFSISLIFLPMCNHVSALASNYLSLMTDDPVLNGTOHTKRLSVYG 1020
Dy 961 SVYDYNKAIQGLFSISLIFLPMCNHVSALASNYLSLMTDDPVLNGTOHTKRLSVYG 1020
Dy 960 SVYDYNKAIQGLFSISLIFLPMCNHVSALASNYLSLMTDDPVLNGTOHTKRLSVYG 1019
Dy 1021 ALGISQGIANFGYSMANISIGIILASRCLHVDLHSLIRSPMSPEERTSGULVNRFSKL 1080
Dy 1021 ALGISQGIANFGYSMANISIGIILASRCLHVDLHSLIRSPMSPEERTSGULVNRFSKL 1080
Dy 1020 ALGISQGIATFGYSMAVISIGIFASRLHDLHNVLRSPISFEKRPISGLVNRFSKL 1079
Dy 1081 DTVDSMTPEVYKMPGSLFNVYIGACIYILATPIAIIIPPLGIYEFVORFYAASSROL 1140
Dy 1080 DTVDSMTPEVYKMPGSLFNVYIGACIYILATPIAIIIPPLGIYEFVORFYAASSROL 1139
Dy 1141 KRLSVSRSPVYSHNETLLGVSVYIRAPBEOERFIHOSDLKVDENOKAYPSIVANTWLA 1200
Dy 1140 KRLSVSRSPVYSHNETLLGVSVYIRAPBEOERFIHOSDLKVDENOKAYPSIVANTWLA 1199
Dy 1201 VRLCEGNCIYLPALFAVISRHSLSAGVLSVSLQVTTYNTLVRSSEMETNIVA 1260
Dy 1200 VRLCEGNCIYLPALFAVISRHSLSAGVLSVSLQVTTYNTLVRSSEMETNIVA 1259
Dy 1261 VERLKEYSETEKEAPMOIOETAPSPSWPQYGRVFRNYCLRYREDLDFVLRIHIVTINGG 1320
Dy 1260 VERLKEYSETEKEAPMOIOETAPSPSWPQYGRVFRNYCLRYREDLDFVLRIHIVTINGG 1319
Dy 1321 EKVGIAGRGTGAKSSLLTGLFRINESAGEIIIDGINIAKIGLHDLRPFKTIIPQDPVL 1380
Dy 1320 EKVGIAGRGTGAKSSLLTGLFRINESAGEIIIDGINIAKIGLHDLRPFKTIIPQDPVL 1379
Dy 1381 SSGSRNULDPSPQYSDERWTSLELAHLKDPVSLPKLHCEKAGEBNLSVGRQVYCL 1440
Dy 1380 SSGSRNULDPSPQYSDERWTSLELAHLKDPVSLPKLHCEKAGEBNLSVGRQVYCL 1439
Dy 1441 ABALIRKTKIIVLDEAPAAVDLETDDLIOSTIRFOFEDCTVLTATHRNTIMDTTRVIVL 1500
Dy 1440 ABALIRKTKIIVLDEAPAAVDLETDDLIOSTIRFOFEDCTVLTATHRNTIMDTTRVIVL 1499
Dy 1501 DKGEIOBYGAPSDLLQORGLFYSMAKDAGLV 1531

QY 977 LSIPLFCMNHVSALASNNWLSMTWDDPIYNGTQHTKRLSYGALGISOIAVFGYSMA 1036
 DB 846 LSIFLFCMNHVSALASNNWLSMTWDDPIYNGTQHTKRLSYGALGISOIAVFGYSMA 905
 QY 1037 VSIGGILASRCALVDLHLSILASPMSPFERPESGMLNVRFSKELTVDSMIPEVKMFMG 1096
 DB 906 VSIGGILASRCALVDLHLSILASPMSPFERPESGMLNVRFSKELTVDSMIPEVKMFMG 965
 QY 1097 SLFNIVGACIVILLATPTIAAIIIPPLGLIFFVQRFYVASSHQRLKLESVSPYVSHN 1156
 DB 966 SLFNIVGACIVILLATPTIAAIIIPPLGLIFFVQRFYVASSHQRLKLESVSPYVSHN 1025
 QY 1157 FTLLGVSTIRAEERERFIHQSDLVKVDENQKAYYSISYANRWLVRLBCVGCIVLPAAL 1216
 DB 1026 FTLLGVSTIRAEERERFIHQSDLVKVDENQKAYYSISYANRWLVRLBCVGCIVLPAAL 1085
 QY 1217 FAVISRHSLSAGLVGSLVSYSLQVTTYLNLVMSSEMETNVAVERLKEVSETEKAPW 1276
 DB 1086 FAVISRHSLSAGLVGSLVSYSLQVTTYLNLVMSSEMETNVAVERLKEVSETEKAPW 1145
 QY 1277 QIOETAPSSWMPQVGRVERFRNYCLARYEDLDVLRHINVTINGEKVGIWRTGAGKSSL 1336
 DB 1146 QIOETAPSSWMPQVGRVERFRNYCLARYEDLDVLRHINVTINGEKVGIWRTGAGKSSL 1205
 QY 1337 TLGLFRINESAGEIIIDGINAKIGLHDLRKITITIPDPPLFSGSLRPMIDPSSQYSD 1396
 DB 1206 TLGLFRINESAGEIIIDGINAKIGLHDLRKITITIPDPPLFSGSLRPMIDPSSQYSD 1265
 QY 1397 EEVWTSLELAHKDVPVSLPDKLDHECAEGENLSVGOROLCLARALLRKTIVLDEA 1456
 DB 1266 EEVWTSLELAHKDVPVSLPDKLDHECAEGENLSVGOROLCLARALLRKTIVLDEA 1325
 QY 1457 TAAVLDLTDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYRVVLVLDKGEIOEGAPSDLLQ 1516
 DB 1326 TAAVLDLTDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYRVVLVLDKGEIOEGAPSDLLQ 1385
 QY 1517 QRGLEYMAKDAGLV 1531
 DB 1386 QRGLEYMAKDAGLV 1400

RA Schriml L.M., Kanapin A., Matsuda H., Betalov S., Beisel K.W.,
 Blake J.A., Bradt D., Brusic V., Chochia C., Corbati L.E., Cousins S.,
 Dalla E., Dregant T.A., Fletcher C.F., Forrest A., Fraser K.S.,
 Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 Kani A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 Konagaya A., Kurochkin I.V., Lee Y., Lehman B., Lyons P.A.,
 Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
 Petrovsky N., Pillat R., Pontius J.U., Qi D., Ramchandran S.,
 Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 Verardo R., Wagner T., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 Wilming L.G., Wysshaw-Bois A., Yangisawa M., Yang I., Yang L.,
 Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carinici P., Hayatsu N.,
 Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
 Hara A., Hashizume W., Imocant K., Ishii Y., Itoh M., Kagawa I.,
 Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 CC -1- FUNCTION: May participate directly in the active transport of
 drugs into subcellular organelles or influence drug distribution
 indirectly (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 DR EMBL: AF022908; AACB0938.1; -; mRNA.
 DR EMBL: AK029876; BAC26654.1; -; mRNA.
 DR HSBP: P08716; IMTO.
 DR Ensembl: ENSMUSG00000023086; Mus musculus.
 DR MGI: MGI:102676; Abcc1.
 DR GO: GO:0005887; C:integral to plasma membrane; IDA.
 DR InterPro: IPR003593; AAA ATPase.
 DR InterPro: IPR011527; ABC_membrane_1.
 DR InterPro: IPR001140; ABC_TM_transp.
 DR InterPro: IPR003439; ABC_transp_1like.
 DR InterPro: IPR005292; MRP_assoc.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PP000006; ABC_transporter; 2.
 DR SMART: SM00382; AAA; 2.
 DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE: PS50929; ABC_TM1F; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE: PS50893; ABC_TRANSPORTER_2; 2.
 KM ATP-binding; Glycoprotein; Nucleotide-binding; Repeat; Transmembrane;
 KM Transport.
 FT TOPO_DOM 1 33 Extracellular (By similarity).
 FT TRANSMEM 34 54 1 (By similarity).
 FT TOPO_DOM 55 74 Cytoplasmic (By similarity).
 FT TRANSMEM 75 95 2 (By similarity).
 FT TOPO_DOM 96 100 Extracellular (By similarity).
 FT TRANSMEM 101 121 3 (By similarity).
 FT TOPO_DOM 122 133 Cytoplasmic (By similarity).
 FT TRANSMEM 134 154 4 (By similarity).
 FT TOPO_DOM 155 172 Extracellular (By similarity).
 FT TRANSMEM 173 193 5 (By similarity).
 FT TOPO_DOM 194 317 Cytoplasmic (By similarity).
 FT TRANSMEM 318 338 6 (By similarity).
 FT TOPO_DOM 339 364 Extracellular (By similarity).

FT TRANSMEM 365 385 7 (By similarity).
FT TOPO DOM 386 441 Cytoplasmic (By similarity).
FT TRANSMEM 442 462 8 (By similarity).
FT TOPO DOM 463 465 Extracellular (By similarity).
FT TRANSMEM 466 486 9 (By similarity).
FT TOPO DOM 487 548 Cytoplasmic (By similarity).
FT TRANSMEM 549 569 10 (By similarity).
FT TOPO DOM 570 591 Extracellular (By similarity).
FT TRANSMEM 592 612 11 (By similarity).
FT TOPO DOM 613 663 Cytoplasmic (By similarity).
FT TRANSMEM 664 984 12 (By similarity).
FT TOPO DOM 985 1022 Extracellular (By similarity).
FT TRANSMEM 1023 1043 13 (By similarity).
FT TOPO DOM 1044 1086 Cytoplasmic (By similarity).
FT TRANSMEM 1087 1107 14 (By similarity).
FT TOPO DOM 1108 1108 Extracellular (By similarity).
FT TRANSMEM 1109 1129 15 (By similarity).
FT TOPO DOM 1130 1200 Cytoplasmic (By similarity).
FT TRANSMEM 1201 1221 16 (By similarity).
FT TOPO DOM 1222 1223 Extracellular (By similarity).
FT TRANSMEM 1224 1244 17 (By similarity).
FT TOPO DOM 1245 1528 Cytoplasmic (By similarity).
FT DOMAIN 326 609 ABC transmembrane type-1.1.
FT DOMAIN 644 868 ABC transporter 1.
FT DOMAIN 971 1253 ABC transmembrane type-1.2.
FT NP_BIND 1290 1524 ABC transporter 2.
FT NP_BIND 678 685 ATP 1 (Potential).
FT NP_BIND 1324 1331 ATP 2 (Potential).
FT CARBOHYD 19 19 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1003 1003 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1528 AA; 171185 MW; 68FD13667D61DBB CRC64;

Query Match 69.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 86.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFSGADSDMLDMNVTWNTSNPDFTCFQNTLVWPCRYLWACPFYFLYSRH 60
DB 1 MALRSPFSGADSDMLDMNVTWNTSNPDFTCFQNTLVWPCRYLWACPFYFLYSRH 60
QY 61 DRGVIQMTPLNKTALGFLMLTYCMADLFYSFWERSGFIPLAFVLSPTLGLITLLA 120
DB 61 DRGVIQMTPLNKTALGFLMLTYCMADLFYSFWERSGFIPLAFVLSPTLGLITLLA 120
QY 121 TPLQLERRRGVSGGIMLTFWLVVCAALIRSKIMTAKEDAOVDFPDITFYVFS 180
DB 121 TPLQLERRRGVSGGIMLTFWLVVCAALIRSKIMTAKEDAOVDFPDITFYVFS 180
QY 181 LLLIQLVSCGSDSPLESTIHDNPPCPRESSASFLRITFWMTGLIVGGRPLEGSD 240
DB 181 LLLIQLVSCGSDSPLESTIHDNPPCPRESSASFLRITFWMTGLIVGGRPLEGSD 240
QY 241 LMSLNKEDTSEQVAVLVKMKKECAKTRKOPVKVYS-SKDPAPKRESSKVDANEVEA 299
DB 241 LMSLNKEDTSEQVAVLVKMKKECAKTRKOPVKVYS-SKDPAPKRESSKVDANEVEA 299
QY 300 LIYKSPKEMNPISLFKLYKTGPYFLMSEFFKAHDLMMESGQIILKLIKFNVDKAP 359
DB 300 LIYKSPKEMNPISLFKLYKTGPYFLMSEFFKAHDLMMESGQIILKLIKFNVDKAP 359
QY 360 DMQGYFTVLLFVACQTLVHOYFHCIFVSGKRITAVIGAVYRKALVITNSAKSST 419
DB 360 DMQGYFTVLLFVACQTLVHOYFHCIFVSGKRITAVIGAVYRKALVITNSAKSST 419
QY 420 VGEIVNLSYDAQRFMDLATYINMIWSAPLOVILATILMLNLSPSYLAGAVVAVLWVP 479
DB 420 VGEIVNLSYDAQRFMDLATYINMIWSAPLOVILATILMLNLSPSYLAGAVVAVLWVP 479
QY 480 NAYVAKTKTYQVAHMSKDNRIKLAMEILNGIVLVKLYAMELAFKQVTAIROEELKVL 539
DB 480 NAYVAKTKTYQVAHMSKDNRIKLAMEILNGIVLVKLYAMELAFKQVTAIROEELKVL 539
QY 540 KKSAYLSAVGTFTVVCIPFLVALCTPAVYVITIDENNLIDQATAVSLAPRIILRPLNLT 599
DB 540 KKSAYLSAVGTFTVVCIPFLVALCTPAVYVITIDENNLIDQATAVSLAPRIILRPLNLT 599

DB 541 KKSAYLSAVGTFTVVCIPFLVALCTPAVYVITIDENNLIDQATAVSLAPRIILRPLNLT 600
QY 600 PMVTSIVQASVSKRLRIFLSHELEPDSIERRPVQGGTNSITVRNATFTWASDDP 659
DB 601 PMVTSIVQASVSKRLRIFLSHELEPDSIERRSIKSGEG-NSITVKNATFTWARGEP 659
QY 660 TUNGTITSIEGALVAVVGVCGCKSSLTALLAEMKVEGHVAKSVAVVPOQATON 719
DB 660 TUNGTITSIEGALVAVVGVCGCKSSLTALLAEMKVEGHVAKSVAVVPOQATON 719
QY 720 DSLRENTLFCQLEEPYRSVITQACALLPDLLEILPSGDRTEIGEKVNLSCGQCORSLA 779
DB 720 DSLRENTLFCQLEEPYRSVITQACALLPDLLEILPSGDRTEIGEKVNLSCGQCORSLA 779
QY 780 RAVYSNADIVLPDDPLSAVDAAHVAKHIFENVIGKMKKTKRILLVTHSMSTLPQVDVI 839
DB 780 RAVYSNADIVLPDDPLSAVDAAHVAKHIFENVIGKMKKTKRILLVTHSMSTLPQVDVI 839
QY 840 VMSGKISEMSGVQELLARQGAFAEFRTYASTFQEDDAENGVTGVSGPKKAKOMENG 899
DB 840 VMSGKISEMSGVQELLARQGAFAEFRTYASTFQEDDAENGVTGVSGPKKAKOMENG 899
QY 900 MLVTSAGKQLOLQSSSSYSVSDISRHNSTAELOKAKKERTWKLMEADKQTCQVK 959
DB 900 MLVTSAGKQLOLQSSSSYSVSDISRHNSTAELOKAKKERTWKLMEADKQTCQVK 959
QY 959 LSVYVNMKAIIGLFTITLIFLCNHNVSALSNYMLSTWTD-PYVNGQHTKRLSV 1018
DB 959 LSVYVNMKAIIGLFTITLIFLCNHNVSALSNYMLSTWTD-PYVNGQHTKRLSV 1018
QY 1019 YGALGISQGIADVRCYMAVSIIGILASRCLHVDLHSLNSPMSFEERTSGNLYNRSK 1078
DB 1019 YGALGISQGIADVRCYMAVSIIGILASRCLHVDLHSLNSPMSFEERTSGNLYNRSK 1078
QY 1078 ELDTVDMSIPEVIMKFMGSLFNVTGACIVILTLATPIAIIIPGLGYFFQRPYVASR 1138
DB 1078 ELDTVDMSIPEVIMKFMGSLFNVTGACIVILTLATPIAIIIPGLGYFFQRPYVASR 1138
QY 1139 QLRLESVSRSPPVSHNETLIGSVIRAFEBEORFIHQSDLKVDENQKAYYPSIVANRW 1198
DB 1139 QLRLESVSRSPPVSHNETLIGSVIRAFEBEORFIHQSDLKVDENQKAYYPSIVANRW 1198
QY 1199 LAVRLECVGNCIVLPAALFVVISRHSLSAGLVGSYSISQVTTYLNMVLRMSSEMTNI 1258
DB 1199 LAVRLECVGNCIVLPAALFVVISRHSLSAGLVGSYSISQVTTYLNMVLRMSSEMTNI 1258
QY 1259 VAVERLKEVSETEKEAPMOIOETRAPSPSQVGRVFRNCLRRREDLDFVLRHINTIN 1318
DB 1259 VAVERLKEVSETEKEAPMOIOETRAPSPSQVGRVFRNCLRRREDLDFVLRHINTIN 1318
QY 1319 GGEKVGIVGRTAGKSSLTGLFRINESAGEIITIDINIAKIGLHDLRFRKTIITIPDPV 1378
DB 1319 GGEKVGIVGRTAGKSSLTGLFRINESAGEIITIDINIAKIGLHDLRFRKTIITIPDPV 1378
QY 1379 LFSGLSRMNLDPSPQYSDEEVMWTSLEIAHLKDFVSALPDKLDHCAEGENLSVGQRLV 1438
DB 1379 LFSGLSRMNLDPSPQYSDEEVMWTSLEIAHLKDFVSALPDKLDHCAEGENLSVGQRLV 1438
QY 1439 CLARALLRKTIIVLEBATAVAVLETDNLQSTIRQFEDCTVLTIAHRLNTIMDYRVI 1498
DB 1439 CLARALLRKTIIVLEBATAVAVLETDNLQSTIRQFEDCTVLTIAHRLNTIMDYRVI 1498
QY 1499 VLDKGEIOEYCAPSDLLQQRGLFYSMAKDAGLV 1531
DB 1499 VLDKGEIOEYCAPSDLLQQRGLFYSMAKDAGLV 1531

RESULT 12
Q810E4 RAT PRELIMINARY; PRT; 1532 AA.
ID Q810E4 RAT
AC Q810E4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ATP-binding cassette protein C1.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Yabuchi H., Takayanagi S., Ishikawa T.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF487549; AA085437.1; -; mRNA.
DR HSSP; P08716; 1MT0.
DR Ensembl; ENSRNOG00000032748; Rattus norvegicus.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity; coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldo/ket_red.
DR Pfam; PF00664; ABC_assoc.
DR Pfam; PF00005; ABC_membrane_2.
DR Pfam; PD00006; Abc_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDO_KETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1532 AA; 171493 MW; 2E6939F63FEA3F68 CRC64;

Query Match 69.2%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;

QY 1 MALRGFCADGSDPLMDNNTWNTSNPDFTKCFONTVLVWPCFYLMACEPFYFLYSRH 60
DB 1 MALSFCCSSDGDPLMDNNTWNTSNPDFTKCFONTVLVWPCFYLMACEPFYFLYSRH 60
QY 61 DRGYQMTPLNKTALGFPLMIVCMADLFYFMRSGIFLAPFLVSPILLGTTTLA 120
DB 61 DRGYQMTPLNKTALGFPLMIVCMADLFYFMRSGIFLAPFLVSPILLGTTTLA 120
QY 121 TFLILERRKGVSSGIMLTFWLVALVCALILRSKIMTALKEADQVDFRDITFYVES 180
DB 121 TFLILERRKGVSSGIMLTFWLVALVCALILRSKIMTALKEADQVDFRDITFYVES 180
QY 121 TFLILERRKGVSSGIMLTFWLVALVCALILRSKIMTALKEADQVDFRDITFYVES 180
DB 121 TFLILERRKGVSSGIMLTFWLVALVCALILRSKIMTALKEADQVDFRDITFYVES 180
QY 181 LILQLVLSCFSDRSPLFSETHDNPCESSASFLSRITFWITGLIVGRQPLESGD 240
DB 181 LILQLVLSCFSDRSPLFSETHDNPCESSASFLSRITFWITGLIVGRQPLESGD 240
QY 181 LILQLVLSCFSDRSPLFSETHDNPCESSASFLSRITFWITGLIVGRQPLESGD 240
DB 181 LILQLVLSCFSDRSPLFSETHDNPCESSASFLSRITFWITGLIVGRQPLESGD 240
QY 241 LMSLNKERTSEOVVLYVNMKKECAKTRKQPVKYVYS-SKDPAPKESKIDANEEVYA 299
DB 241 LMSLNKERTSEOVVLYVNMKKECAKTRKQPVKYVYS-SKDPAPKESKIDANEEVYA 299
QY 241 LMSLNKERTSEOVVLYVNMKKECAKTRKQPVKYVYS-SKDPAPKESKIDANEEVYA 299
DB 241 LMSLNKERTSEOVVLYVNMKKECAKTRKQPVKYVYS-SKDPAPKESKIDANEEVYA 299
QY 300 LIVKSPKEMNPSLKFVLYKTFGPFYFMSFFPKAHLDMFSGPOLKLLIKFVNDTKAP 359
DB 301 LIVKSPKEMNPSLKFVLYKTFGPFYFMSFFPKAHLDMFSGPOLKLLIKFVNDTKAP 359
QY 360 DWOGFYTVLLFVTACLQTLVLAHQYFHI CFVSGMRIKTAVIGAVYRKALVITNSARKSST 419
DB 360 DWOGFYTVLLFVTACLQTLVLAHQYFHI CFVSGMRIKTAVIGAVYRKALVITNSARKSST 419

DB 361 DWOGFYTVLLFVSACIQTLALHQYFHI CFVTGNRIKTAVIGAVYRKALVITNSARKSST 420
QY 420 VGEIYVNLMSYDAQFMFLATYINMISAPLOVIALYILMLNLGSPYLAGVAVVWLMPV 479
DB 421 VGEIYVNLMSYDAQFMFLATYINMISAPLOVIALYILMLNLGSPYLAGVAVVWLMPV 480
QY 480 NAYVAMKTKTYQVAMHMSKDNRIKLMNEILNGIKYLVKLYAMELAFKDKVLAIRQELKVL 539
DB 481 NAYVAMKTKTYQVAMHMSKDNRIKLMNEILNGIKYLVKLYAMELAFKDKVLAIRQELKVL 540
QY 540 KKSAYLVAVGTFPTVWCCTPELVALCTPAVYVYTDENNILDAQTAVASLAPFIIAEPNLT 599
DB 541 KKSAYLVAVGTFPTVWCCTPELVALCTPAVYVYTDENNILDAQTAVASLAPFIIAEPNLT 600
QY 600 PMVTSIYQSVSLKRLRIFLSHELEPDSIERPVPKDGSGTNSITYRANFTYARSDP 659
DB 601 PMVTSIYQSVSLKRLRIFLSHELEPDSIERPVPKDGSGTNSITYRANFTYARSDP 660
QY 660 TLNGITFSIPBGLVAVVGVGCGKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAMION 719
DB 661 TLNGITFSIPBGLVAVVGVGCGKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAMION 720
QY 720 DLSAENILFGQLEBPYRYSTYQACALLPLELTPSGDRMEIGKGVNLGGGQKORUSLA 779
DB 721 DLSAENILFGQLEBPYRYSTYQACALLPLELTPSGDRMEIGKGVNLGGGQKORUSLA 780
QY 780 RAYVSNADYIFDDPLSAVDHVGKHI FENYIGPKMKLNKTRILVTHSMGYLPQVDYII 839
DB 781 RAYVSNADYIFDDPLSAVDHVGKHI FENYIGPKMKLNKTRILVTHSMGYLPQVDYII 840
QY 840 VMSGGKISEMGSYQELLARDGAPAEPLRTYASTQEOBDAEENGTVGSGPKKAKOMENG 899
DB 841 VMSGGKISEMGSYQELLARDGAPAEPLRTYASTQEOBDAEENGTVGSGPKKAKOMENG 900
QY 900 MLVYDSAGKQLOQRLSSSSSYSGDISSRHNSHTALQRAEAKKETWTLMEADKQOTGVK 959
DB 901 MLVYDSAGKQLOQRLSSSSSYSGDISSRHNSHTALQRAEAKKETWTLMEADKQOTGVK 960
QY 960 LSVYVDYKAKIGLFIISFIFLPMCNHVSALASNYWLSMTDD-PYVNGTOEHTKVRSLV 1018
DB 961 LSVYVDYKAKIGLFIISFIFLPMCNHVSALASNYWLSMTDD-PYVNGTOEHTKVRSLV 1019
QY 1019 YGALGISQGLAVFYSMAVSIIGGILASRCLHVDLLHSILNSPMSFFERTPGSNLVNRSK 1078
DB 1020 YGALGISQGLAVFYSMAVSIIGGILASRCLHVDLLHSILNSPMSFFERTPGSNLVNRSK 1079
QY 1079 ELDTYDSMTPEVYIKMFSGSLFNVTGACVILALPPIAIIIPPLGLIYFFQRYPVASSR 1138
DB 1080 ELDTYDSMTPEVYIKMFSGSLFNVTGACVILALPPIAIIIPPLGLIYFFQRYPVASSR 1139
QY 1139 QKRLLESYSRSPVYSHFNETLLGVSITRAPFEOERFIHQSLKXVDENOKAYYPSIVANRW 1198
DB 1140 QKRLLESYSRSPVYSHFNETLLGVSITRAPFEOERFIHQSLKXVDENOKAYYPSIVANRW 1199
QY 1199 LAVRLAEVGNCTIVPALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNI 1258
DB 1200 LAVRLAEVGNCTIVPALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNI 1259
QY 1259 VAVRLKAYSTETKEAPWQIOETAPPSWPQOVGKVEFPNNTYCLARTREDDLPRLHINVTIN 1318
DB 1260 VAVRLKAYSTETKEAPWQIOETAPPSWPQOVGKVEFPNNTYCLARTREDDLPRLHINVTIN 1319
QY 1319 GGEKVYIGRTGAGKSSLTGLFPIINSABEIIIDGINIKIGLHDLRFKITTIPQDPV 1378
DB 1320 GGEKVYIGRTGAGKSSLTGLFPIINSABEIIIDGINIKIGLHDLRFKITTIPQDPV 1379
QY 1379 LFSGSLRKNLDPFQOYSDDEEYWTSLAHLADFYVALPDKLDHCAEGENLVSQOROLV 1438
DB 1380 LFSGSLRKNLDPFQOYSDDEEYWTSLAHLADFYVALPDKLDHCAEGENLVSQOROLV 1439
QY 1439 CLARALLKTKIILVDEATAVADLETDDLIOSTRTOFQEDCTVTIARLNTIMDYTVI 1498
DB 1440 CLARALLKTKIILVDEATAVADLETDDLIOSTRTOFQEDCTVTIARLNTIMDYTVI 1499


```
QY      1499  VLDKGEIQEYGAPSDLLQORGLFYSMAKDAGLV 1531
      |||||:||||:||||:||||:||||:||||:||||:
DB      1500  VLDKGEIIRECGAPSELLQORGFYSMAKDAGLV 1532
```

RESULT 13

Q8CG09_RAT	
ID	Q8CG09_RAT
	PRELIMINARY;
	PRT;
	1532 AA.

	DT	01-MAR-2003 (TREMBlrel. 23, Created)
	DT	01-MAR-2003 (TREMBlrel. 23, Last sequence update)
	DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)
	DE	Multidrug resistance-associated protein 1.
	GN	Name=Abcc1; Synonyms=Mrp1;
	OS	Rattus norvegicus (Rat).
	OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
	OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
	OC	Muridae; Murinae; Rattus.
	OX	NCBI_TaxID=10116;
	RN	[1]
	RP	NUCLEOTIDE SEQUENCE.
	RC	STRAIN=Sprague-Dawley; TISSUE=Brain;
	RA	Yang Z., Li C.S.W., Shen D.D., Ho R.U.Y.;
	RL	Submitted (Oct-2002) to the EMBL/genbank/DBJ databases.
	CC	-1- SIMILARITY: Belongs to the ABC transporter family.
	DR	EMBL; AY170916; AAN86532.1; -. mRNA.
	DR	HSSP; P08716; IMTO.
	DR	Ensembl; ENSRNCG0000032748; Rattus norvegicus.
	DR	RGD; 3112; Abcc1.
	DR	GO; GO:0016021; C:integral to membrane; IEA.
	DR	GO; GO:0005524; F:ATP binding; IEA.
	DR	GO; GO:0016887; F:ATPase activity; IEA.
	DR	GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
	DR	GO; GO:0000160; F:nucleotide binding; IEA.
	DR	GO; GO:0006810; F:transport; IEA.
	DR	InterPro; IPR003593; AAA_ATPase.
	DR	InterPro; IPR011527; ABC_membrane_1.
	DR	InterPro; IPR001140; ABC_TM_transp.
	DR	InterPro; IPR003439; ABC_tranp_like.
	DR	InterPro; IPR001395; Aldo/ket_Fed.
	DR	InterPro; IPR005292; MRP_assoc.
	DR	Pfam; PF00664; ABC_membrane; 2.
	DR	Pfam; PF00005; ABC_tran; 2.
	DR	ProDom; PD000006; ABC_transporter; 2.
	DR	SMART; SMO0382; AAA; 2.
	DR	TIGRFAMS; TIGR00957; MRP_assoc_pro; 1.
	DR	PROSITE; PS00929; ABC_TMIF; 2.
	DR	PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
	DR	PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
	DR	PROSITE; PS00063; ALDO-KETO REDUCTASE 3; UNKNOWN_1.
	KW	ATP-binding; Nucleotide-binding; Repeat.
	SQ	SEQUENCE 1532 AA; 171558 MW; C18F3534DD85732A CRC64;
		Query Match 69.2%; Score 6932.5; DB 2; Length 1532;
		Best Local Similarity 87.3%; Pred. No. 0;
		Matches 1338; Conservative 100; Mismatches 92; Indels 3; Gaps 3
QY	1	MALGFGCAGDSDDLMDNNVTMTNSNDPTKCFONTLVVWPCFYIACFPFYLYLSRH 60
DB	1	MAIKSFSCSDSDSLMDNNVTMTNSNDPTKCFQNTVLTVWPCFLYMSCFPLXLYLSRH 60
QY	61	DGGYIQMPELKKTKALGFLLMIYVMADLFYSFWERSRGFIAPAFIVLSPTLLGTTLIA 120
DB	61	DGGYIQMTHLNKAKTALGFLLMIYVMADLFYSFWERSQGLAPVLVLSPTLLGTTLIA 120
QY	121	TFLIQLERRKGVSQSGIMLTFMLVALVCALAIIRSKINTALKEDAQVDLFRDIIFYVES 180
DB	121	TFLIQLERRKGVSQSGIMLTFMLVALVCALAIIRSKISIALKQDAQVDMFRDSAYLYVFT 180
QY	181	LILLIQLVLTSCGSDSPLESETIHDPNCPBESSASTLSRTITTWITGLIVRGYRQPLES 240
DB	181	LVIQLVLTSCSDSPLESEIVVRPNCPBESSASTLSRTITTWITGMVVGYPQLKS 240

QY	241	LMSLNKEPTSSQVVPVLVNNWKKCBKACTRKQRPVKVYS-SKQPAQPKSSKYDANEVBA	299
Db	241	LMSLNKEPTSSQVVPVLVNNWKKCBKCVRSRKQPVRIVAPRQDPTKPKSSQOLDVNEEVEBA	300
QY	300	LIVSPKQENPNSL.FKVL.YKT.FGYP.LMS.FE.FFK.IHDI.MMS.GPO.I.LK.LI.KF.VND.TAP	359
Db	301	LIVSSHDRPSPSL.FKVL.YKT.FGEP.LMS.FY.KAL.HDI.MF.AGP.IE.LI.IIN.FV.NDR.AP	360
QY	360	DMQGYFYVLL.FVYI.ACTQTLVIAQY.FHI.CFVSG.MRI.KTAAV.GAVRKAL.VITNS.ARK.SST	419
Db	361	DMQGYIYALL.FVSA.CI.QTLALH.QY.FHI.CFYTG.RIKTAAV.GAYYRKAL.VITNS.ARK.SST	420
QY	420	VGEIYVNLMSYDAQR.FMDLATY.I.NMI.WSAP.LQV.TIAL.YFL.MLNL.GBSVLAGYA.VMYLAMPV	479
Db	421	VGEIYVNLMSYDAQR.FMDLATY.I.NMI.WSAP.LQV.TIAL.YFL.MLNL.GBSVLAGYA.VMYLAMPV	480
QY	480	NAVMAAKTKTYQVAHMKSKNRIKLMNFI.LNGIKYV.KUYAMELAFQDKVLA.IROBEL.KVL	539
Db	481	NAVMAAKTKTYQVAHMKSKNRIKLMNFI.LNGIKYV.KUYAMELAFQDKVLA.IROBEL.KVL	540
QY	540	KKSAYLSAVGFTWV.CPFLV.ALC.PAYVYV.TIDEN.NI.DAQ.TAF.SI.L.FN.I.RP.LI.II	599
Db	541	KKSAYLSAVGFTWV.CPFLV.ALS.TFAV.FV.DEN.NI.DAKA.FV.SI.L.FN.I.RP.LI.II	600
QY	600	PMVISTIVQASVSLKRLRI.FLSHELEB.DS.IERR.PYKDG.GG.TNS.ITV.NA.NFTW.ARSDP	659
Db	601	PMVISTIVQASVSLKRLRI.FLSHELEB.DS.IERRS.IKDG.GM.NS.ITV.NA.NFTW.ARSDP	660
QY	660	TLNGITTSIPRGALVAVVGQVCGCKSSILSLALBMDKVEGHVAIKGSVAVYPOQAWION	719
Db	661	TLNGITTFPGRGALVAVVGQVCGCKSSILSLALBMDKVEGHVTLKGSVAVYPOQAWION	720
QY	720	DSLSENLILFGQLEBPYRYSVIOK.CALLPDL.EI.LP.SG.RT.IEGE.GVNL.SSGQQRVS.LA	779
Db	721	DSLSENLILFGPRLQEBCHKAWME.CALLPDL.EI.LP.SGGLTE.IEGE.GVNL.SGQQRVS.LA	780
QY	780	RAVYSNADIY.FPDDP.DLSAVDAHVGKHI.FENYIGPKMKONTRLI.VTHSMVS.YL.POV.DY.I	839
Db	781	RAVYCNSDIY.LDDP.DLSAVDAHVGKHI.FEKVYVGMGLAKNTRLI.VTHGIS.YL.POV.DY.I	840
QY	840	VMSGGKISSEMGSYQELLARDGAFAPLRTYVASTEOQDABENGVTGVS.GPKGEAKOMENG	899
Db	841	VMSGGKISSEMGSYQELLARDGAFAPLRTYVANTQDLASEDSSKMGVSGLCKESKPENG	900
QY	900	MLVYDSAGKQLOQOROSSSSSYSGDI.SKRHNSTAI.LQRAKAKETWKLMEADKAQOTGVK	959
Db	901	ILVYDAVGKPLQORHLSNSSSHSVVTTNOHSTAEI.LQKS-GYKEBETWKLMEADKAQOTGVK	959
QY	960	LSVYMDWTKAIGLISFLSIFL.FMCHNYSALASVYMLSLMTDD-PYNGTOEHKRVRLSV	1018
Db	960	LSVYMDWTKAIGLISFLSIFL.FMCHNYSALASVYMLSLMTDDP.PYAVNGTOENNRFLSV	1019
QY	1019	YGALGISQGLIAVFYSMAVSI.GGILARCLHVDLHLSILRS.PMS.FE.FKTPSGN.LVNR.FSK	1078
Db	1020	YGALGISQGLIAVAFYSMAVSI.GGILARCLHVDLHLSILRS.PMS.FE.FKTPSGN.LVNR.FSK	1079
QY	1079	ELDTVDNMIPEVIMFMGSLFNVI.GACTVILATPIA.II.IIPGLIYFVQRFVASSR	1138
Db	1080	ELDTVDNMIPEVIMFMGSLFNVI.GACTVILATPIA.II.IIPGLIYFVQRFVASSR	1139
QY	1139	QKRLBESVNS.PVYVSHENFTLIGSVIRAP.FEQR.FTHOSL.KYDENOKA.YYPSI.VANRW	1198
Db	1140	QKRLBESVNS.PVYVSHENFTLIGSVIRAP.FEQR.FTHOSL.KYDENOKA.YYPSI.VANRW	1199
QY	1199	LAVRLECVGNCQIV.FAAL.FAVISRHSLSI.SAGLVGVSYSIQVTTYL.NMLV.MS.SMET.NI	1258
Db	1200	LAVRLECVGNCQIV.FAAL.FAVISRHSLSI.SAGLVGVSYSIQVTTYL.NMLV.MS.SMET.NI	1259
QY	1359	VAVRLEKYSYTEKEAPMOIOETAP.PSSWPQVGEVBERFNTCLARYREDLDFVLRHINVTIN	1318
Db	1360	VAVRLEKYSYTEKEAPMOIOETAP.PSTWPSGSGVEERDQCLARYREDLDFVLRHINVTIE	1319

QY 1319 GGEKVGIVGRTAGKSSLTGLFRINESAGEBIIIDGINIAKIGHLDIRPKITIIIPQDPV 1378
DB 1320 GGEKVGIVGRTAGKSSLTGLFRINESAGEBIIIDGINIAKIGHLDIRPKITIIIPQDPV 1379
QY 1379 LFGSSLRNNLDPFGQYSDREFTWTSLELAHDKDFVSALPDKLDHECAGEGNSVQGRQV 1438
DB 1380 LFGSSLRNNLDPFGQYSDREFTWTSLELAHDKDFVSALPDKLDHECAGEGNSVQGRQV 1439
QY 1439 CLARALLRKTILVLEDEATAVDELTDLLIOSTRIRQFEDCTVLTIARLNTIMDYTRVI 1498
DB 1440 CLARALLRKTILVLEDEATAVDELTDLLIOSTRIRQFEDCTVLTIARLNTIMDYTRVI 1499
QY 1499 VLDKGEIQEGAPSDLLQQRGLFYSMAQAGLV 1531
DB 1500 VLDKGEIQEGAPSDLLQQRGLFYSMAQAGLV 1532
RESULT 14
Q810G9 RAT PRELIMINARY; PRT; 1523 AA.
ID Q810G9 RAT PRELIMINARY; PRT; 1523 AA.
AC Q810G9;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE ATP-binding cassette protein C1 variant A.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Takayanagi S., Ishikawa T.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY174892; AA044983.1; -, mRNA.
DR HSSP; P08716; IMT0.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR01527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_tranp_like.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran_2.
DR ProDom; PD000006; ABC_transporter_2.
DR SMART; SM00382; AAA_2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro_1.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro_1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR ATP-binding; Nucleotide-binding; Repeat.
KW SEQUENCE 1523 AA; 170505 MW; E40337051A1CB9C6 CRC64;
SQ
Query Match 68.8%; Score 6892; DB 2; Length 1523;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;

DB 61 DRGYIOWTHLNKATLTAAGFLMWICWADLFYSFWERSRGIFLAPVFLVSPTLGITTLA 120
QY 121 TFLQLERRKGVSSGIMLTFMLVALVCALALISKIMLTALKEADQVDFRDIFFYYVS 180
DB 121 TFLQLERRKGVSSGIMLTFMLVALVCALALISKIMLTALKEADQVDFRDIFFYYVS 180
QY 181 LLLIQLVLSGFSDBSPFSETIHDNPNCPRESSASFLRITFMWITGLVIRGQROPLEGSD 240
DB 181 LVIQLVLSGFSDBSPFSETIHDNPNCPRESSASFLRITFMWITGLVIRGQROPLEGSD 240
QY 241 LMSLNKEDTSEQVFPVLYKMKKECATRKQPPVYVS-SKDPAPQKSSKVDANEVEA 299
DB 241 LMSLNKEDTSEQVFPVLYKMKKECATRKQPPVYVS-SKDPAPQKSSKVDANEVEA 299
QY 300 LVIKSPKEMNPSSLFKVLYKTGEGYPLMSFFFKAIHDLMMSPGQILKLFKFNVDTKAP 359
DB 301 LVIKSHKDRDPSLFKVLKTFGYPFLMSFLYKALHDLMPAGEPILFELINFNVDREAP 360
QY 360 DMQGYFTVLLFVTACIQTIVLAHQYFHCVSNGRKITPAVIGAYRKALVITNSARKSST 419
DB 361 DMQGYLVTLALFVSACIQTIVLAHQYFHCVTGKRITPAVIGAYRKALVITNSARKSST 420
QY 420 VGEIVNLMSVDQAQFPMDLATTYNNMNSAPLOYIALYLMNLGSPVLAGVAVWLAMPV 479
DB 421 VGEIVNLMSVDQAQFPMDLATTYNNMNSAPLOYIALYLMNLGSPVLAGVAVWLAMPV 480
QY 480 NAWMAKTKTYQVAMHMSKDKRIKLMNELLNGIKVLYKLYAMELAFKDVLAIROBELVYL 539
DB 481 NAWMAKTKTYQVAMHMSKDKRIKLMNELLNGIKVLYKLYAMELAFKDVLAIROBELVYL 540
QY 540 KKSAYLSAVGTFTWCTPFLVALCTPAVYVITIDENNILLDAQAFVSLAFNILLRPNLIL 599
DB 541 KKSAYLSAVGTFTWCTPFLVALCTPAVYVITIDENNILLDAQAFVSLAFNILLRPNLIL 600
QY 600 PMVTSIYQASVSLKRLRIFLSHELEPDSIERPPVQDGGTNSITTRNAFTTARSSPP 659
DB 601 PMVTSIYQASVSLKRLRIFLSHELEPDSIERMSIKDGGGMSITVGNATFTARSDPP 660
QY 660 TLNGITTSIPGALVAVVGVGGKSSLLSLALEMDKVEGVAMIKGSVAVPOOAMQN 719
DB 661 TLNGITTSIPGALVAVVGVGGKSSLLSLALEMDKVEGVAMIKGSVAVPOOAMQN 720
QY 720 DSLRENILFGQLEPEPYRSYVQAACALLPDLIELPSGDRIEIGKGVNLSGGQKORVSLA 779
DB 721 DSLRENILFGRLDHECHKAMENACALLPDLIELPSGDRIEIGKGVNLSGGQKORVSLA 780
QY 780 RAVYSNADYIFDPPLSAVDAHVGKHFENYIGPKGMKKNKTRILLVTHSMGYLPQVDYII 839
DB 781 RAVYCNDSIYLLDDPLSAVDHVGKHFEEKVVGEMGLKKNKTRILLVTHGISYLPQVDYII 840
QY 840 VMSGKTSSEMSYQELLARQAFAPLRTYVSTBOBDAEENGYGVSGPKKAKOMENG 899
DB 841 VMSGKTSSEMSYQELLARQAFAPLRTYVSTBOBDAEENGYGVSGPKKAKOMENG 900
QY 900 MLVYDSAGKQOROLSSSSYSYGDISRHHNSTAELQKAEKKEBTWKLMEADKAQTQGVK 959
DB 901 MLVYDSAGKQOROLSSSSYSYGDISRHHNSTAELQKAEKKEBTWKLMEADKAQTQGVK 959
QY 959 LSVYWDYKKAIGLISFLISFLFMCNHSVALASNYLSLWTD-PIYNGTQEHKTRLSV 1018
DB 951 LSVYWNWYKKAIGLISFLISFLFMCNHSVALASNYLSLWTD-PIYNGTQEHKTRLSV 1018
QY 1019 YGALGISGGLAVFGYSMAVSTGGLARSLCHVDLHSLTRSPMSFPERTPSGNLVNRSK 1078
DB 1011 YGALGISGGLAVFGYSMAVSTGGLARSLCHVDLHSLTRSPMSFPERTPSGNLVNRSK 1078
QY 1079 ELDFTVDSMIFEVIKMFMSLFFNVIGACVILATPIAAIIPPLGLIYFFQRFVASSR 1138
DB 1071 ELDFTVDSMIFEVIKMFMSLFFNVIGACVILATPIAAIIPPLGLIYFFQRFVASSR 1138
QY 1139 QLKRLBSVRSRPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENKAYPSIVANRW 1198

Db	1131	QKRLAESVSRPVS	SHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYVPIVANRW	1199
Qy	1199	LAVNLECTGNCITV	PAALFAVIRSRSLGAGVGLSVSLQVTTYLMVLVMSSEMETNI	1258
Db	1191	LAVRECEGCNCITV	FAALFAVISRSHLSAGLVGLSVSLQITTYLMVLVMSSEMETNI	1250
Qy	1259	VAVELRLKYSYSETEKAPQ	IOETAPSPSSMPQVGRVPEPNYCLARYREDLDPLRHNNTIN	1318
Db	1251	VAVERLKYSETETENASQ	IOETAPSPSPSGRVPEPNYCLARRREDLDVLKHNNTIE	1310
Qy	1319	GGEKVGIVGRTGAGKSS	LTLLGLFRINSEAGEIITDGINIAKIGLHDLRFKITTIIPODPV	1378
Db	1311	GGEKVGIVGRTGAGKSS	LTLLGLFRINSEABEIIIDGINIAKIGLHDLRFKITTIIPODPV	1370
Qy	1379	LFSSGLRNLDPFQSY	SDEBWTSLIELAHLKDFYALPDKLDEHCAEGENLSVGQRLV	1438
Db	1371	LFSSGLRNLDPFQSY	SDEBWTSLIELAHLKGFYALPDKLHNECAEGENLSVGQRLV	1430
Qy	1439	CLALALRKTKTILV	DEATAVVDLTDLLIOSTRTOPEDCVTLIARLNTIMDYRVI	1498
Db	1431	CLARLARKTKTILV	DEATAVVDLTDLLIOSTRTOPEDSVTLIARLNTIMDYRVI	1490
Qy	1499	VLDKGEIOEYGAPSD	LLOQRGLFYFSMAKDAGLV	1531
Db	1491	VLDKGEIRCGAPSEL	LQQRGVFFSMADAGLV	1523
RESULT 15				
OSF364	CHICK	OSF364	CHICK PRELIMINARY;	PRT; 1525 AA.
ID	Q5F364	T	10-MAY-2005 (TREMBLrel. 30, Created)	
AC	Q5F364	T	10-MAY-2005 (TREMBLrel. 30, Last sequence update)	
DT	10-MAY-2005	(TREMBLrel. 30, Last annotation update)		
DE	Hypothetical protein.			
GN	ORFNames=RCJMB04.32420;			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxId=9031;			
RL	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=CB; TISSUE=Bursa;			
RA	Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,			
RA	Fiedler L.P., Kuter S., Blagoderstki A., Kostovska D., Kotec M.,			
RA	Plachy J., Garninci P., Hayaishiaki Y., Buerstedde J.M.;			
RT	"Full-length cDNAs from chicken bursa lymphocytes to facilitate			
RT	gene function analysis."			
CC	Genome Biol. 6:R6-R6 (2005).			
CC	-1 - SIMILARITY: Belongs to the ABC transporter family.			
DR	EMBL; AJ851786; CAH65420.1; -; mRNA.			
DR	GO; GO:0016021; C: integral to membrane; IEA.			
DR	GO; GO:0005524; F: ATP binding; IEA.			
DR	GO; GO:0016887; F: ATPase activity; IEA.			
DR	GO; GO:0042626; F: ATPase activity, coupled to transmembrane m. .; IEA.			
DR	GO; GO:0000166; F: nucleotide binding; IEA.			
DR	GO; GO:0006810; P: transport; IEA.			
DR	InterPro; IPR003593; AAA_ATPase.			
DR	InterPro; IPR011527; ABC_membrane_1.			
DR	InterPro; IPR001140; ABC_TM_transp.			
DR	InterPro; IPR003439; ABC_transp_like.			
DR	InterPro; IPR005292; MRP_assoc.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	Pfam; PF00664; ABC_membrane_2.			
DR	Pfam; PF00005; ABC_cran_2.			
DR	ProDom; PD000006; ABC_transporter_2.			
DR	SMART; SM00382; AAA_2.			
DR	TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.			
DR	PROSITE; PS50929; ABC_TM1F; 2.			
DR	PROSITE; PS00211; ABC_TRANSPORTER_1; 2.			
DR	PROSITE; PS50893; ABC_TRANSPORTER_2; 2.			
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN 1.			

KW	ATP-binding, Hypothetical protein; Nucleotide-binding; Repeat;	Seq	Sequence	1525 AA;	170972 MF;	387246B711B878FD CRC64;
Query Match	62.9%; Score 6301; DB 2; Length 1525;					
Best Local Similarity	77.6%; Pred. No. 0;					
Matches 1193;	Conservative 176; Mismatches 148; Indels 20; Gaps 7					
QY	1	MALRGFSAGDS	DPLMDMNTWNTSNPDFTYKCFONTVLVWPCFYLMACPFYLYLSRH	60		
DB	1	MGISTCSADSBEPFMDNLTMHTENPDFTOCFQNTVLVWPCILYWCFPAYFLYLSRH	60			
QY	61	DRGVQMTPLNKTALAGFLMIYCMADLFPSFERSGCIPLAPFLVSPITLGGTITLLA	120			
DB	61	DRGVIQMSLNKATKATAGLLIMICYMADLFPSFERSQNIIRAPFLISPYLGIITMLLA	120			
QY	121	TFLLQLERKKGVSQSGIMLTFMLVALYCALMLISKIMTALKEDQVDLFRDITFYVFS	180			
DB	121	TFLLQHRKKGVSQSGIMMIFMLISLCAIYIPRSKIMLANTDREVAFRVYTCYFI	180			
QY	181	LLLIQLVLSCFSDSPLESETIHDNPNCPRESSASFLSRITFWITGLIYRGYRQPLEGSD	240			
DB	181	LLLVQLISCPREKPIFSEAVNDPKCPESASFLSRITFWITGLIMIQHRRPLEAKD	240			
QY	241	LMSLNKEDTSSQVVPVLVYKMKECATRKQPKVYVYSSKDPAPCKESSKYDAN---	296			
DB	241	LMSLNKEDTSEIIPGLAKMKAMKEMATKRPPLMLLYSSK-----KQKSSDSNGEVMEE	295			
QY	297	VEALIVSPQENKPSLFXLYKTPGPEPLMSFFPKAHIDLMFMSGPOLIKLLIKFVNDT	356			
DB	296	AEALIIYPSQSSSASLSKLYKTFGPEPLMSFKAHIDLMTFGPEILKLLINFWANK	355			
QY	357	KAPDMQGYFYVYVLLFVTACLOTILVLAHYFHLCFPSGMRKITAIVGAVERKALVITNSARK	416			
DB	356	SAPWQGYFTYGLLPVCACLOTILHGYFHLCTYGRMLKTAIVGIVRYKALVITNSARK	415			
QY	417	SSVYGEIVNIMSVDAQRPMDIATYINNIWSAPLOVILLAYLLMLNGBSVLAGAVMVLM	476			
DB	416	TSVYGEIVNIMSVDAQRPMDIATYINNIWSAPLOVILLAYLLMLNGBSVLAGAVMVLL	475			
QY	477	VPNVANAAMTKTYQVAHMKSKDRIKLMENINGIYVLKLYARELA PKDKVYLAIRQSEL	536			
DB	476	VPINAVAMTKTYQVAOMKSKDRIKLMENINGIKYLYARELAEREVLEIRQSEL	535			
QY	537	KVLKKSAYLSAVGFTYVCTPFLVAACTFAVYVITDENINIDAOATAFSLALFNILRPL	596			
DB	536	KVLKKSAYLAAAGFTYVCAFLVAACTFAVYVYVKNINIDAKAFSLALFNILRPL	595			
QY	597	NILPMVITSSIVQASVSLKRLRILFSLSELEBDSIERRPVKQGGTNSITVNNATFTWARS	656			
DB	596	NILPMVITSSIVQASVSLKRLRFLFSHELEBDSIIRGPITVAEG--SIVVKNATFSMKT	653			
QY	657	DPEPLNGITSPRGALVAVYGVQVCGCKSSILSLALBMDKVEGHVAIKGSVAYVPOQAM	716			
DB	654	DPEPLNSINFTVPGSLIAVVGQVCGCKSSILSLALBMDKREGVYVVKGSIAVYPOQAM	713			
QY	717	IQNDLSARENLIFFGQLEBPYRASVIOACALLPDELIIPSGDRTEIGEGVNLSGGQKORV	776			
DB	714	IQNATLEDNIIIFGRENESRYKRYIAECALLPDELIIPMGDRTEIGEGVNLSGGQKORV	773			
QY	777	SLARAAYSNAIDIYFDDPLSAVDAAHVKHILENVIIGPKMLKNKTRILIVTHSMSTLPQVD	836			
DB	774	SLARAAYCNADTYLFDPLSAVDAAHVKHIFEKYIGPKILKNKTRIVLVTHAANYLTPQMD	833			
QY	837	VIIYMSGGKISSENGSYOELLARDCAPLEPLTASTROBDAENGVTGVSGP--GKEAKQ	895			
DB	834	TIIVMTDGEISENGSYOELLAKODAPLEPLRTAYANAQSMESSD-----ASPFGKEKP	888			
QY	896	MENGMVTTDAGKQLOLORLSSSSSSGSDI--SRHNSHTAELQKAEAKKEFTWKMEADKA	953			
DB	889	VENGVAVYNDATGKLMHRQLSNSTRYSKFTGSKGQSSTAEIQRKLAER--NSKULTEADTA	947			
QY	954	QTQGVKLSVYVWDYKALGFLISFLSIFLMCNHNSALASNYWLSLWTDPIVNGTOERTK	1013			

```

Db      948  KTRGVKATVWEYKMAIGLYISFLSVFLMCMNHASLASNYWLSLMTDDPVVNGTQOYTN 1007
Qy      1014  VRLSVYGAIGISOGIAVFGYSMAVSIIGGILASRCHVDLHLSILRSPMSFFERTPSGNLV 1073
Db      1008  VRLGVYGAIGISOGIAVFGYSMAVSIIGGILASRCHVDLHLSILRSPMSFFERTPSGNLV 1067
Qy      1074  NRSFKELDTVDSMIPEVIMKPMGSLFNVYIACIVILATPIAIIIPPLGLYFFVQRFY 1133
Db      1068  SRFSEKIDTIDSTIPPIIKMFSGTFENVYIGACIIILLATPIAAVVIIPPLGLVYLLVQRFY 1127
Qy      1134  VASSRQLRLESVSRSPIYSHNETLLGVSVIRAFEBQERFIHQSDLVYDENOKAYYPSI 1193
Db      1128  VATSRQLRLESVSRSPIYSHNETLLGVSVIRAFEBQERFIHQSDLVYDENOKAYYPSI 1187
Qy      1194  VANRMLAVRLECVGNCIVLFAALFAVISRHSLSAGLVGASVSYSLQVTTYLNWLVYRMSSE 1253
Db      1188  VANRMLAVRLECVGNCIVLFAALFAVISRHSLSAGLVGASVSYSLQVTTYLNWLVYRMSSE 1247
Qy      1254  METNIVAVRLEKESYETEKEAPMOIQETAPPSMPOVGRVREFRNYCLRYREDLDFVLRI 1313
Db      1248  LETNIVAVRLEKESYETEKEAPMOIQETAPPSMPOVGRVREFRNYCLRYREDLDFVLRI 1307
Qy      1314  NVTINGEKVGI VERTGAGKSLTGLPRINSAEGEIIIDGINIAKIGLHDLRPKITII 1373
Db      1308  NVTINGEKVGI VERTGAGKSLTGLPRINSAEGEIIIDGINIAKIGLHDLRPKITII 1367
Qy      1374  PODPVLFGSLRMLNDPFSOYSDEBEWTSLELAHLKDFVSALPDKLDHECAEGENLSVG 1433
Db      1368  PODPVLFGSLRMLNDPFSOYSDEBEWTSLELAHLKDFVSALPDKLDHECAEGENLSVG 1427
Qy      1434  QROLVCLARALIRKTKIIVLDEATAVLDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMD 1493
Db      1428  QROLVCLARALIRKTKIIVLDEATAVLDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMD 1487
Qy      1494  YTRVIVLDRGEVCDSPDNLQAKGLFYSMKXDSGL 1530
Db      1488  YTRVIVLDRGEVCDSPDNLQAKGLFYSMKXDSGL 1524

```

Search completed: December 15, 2005, 15:33:58
 Job time : 260.777 secs